

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:04:25 ; Search time 45.87 Seconds
(without alignments)
52.370 Million cell updates/sec

Title: US-09-701-623c-84

Perfect score: 135

Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 71:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

pred, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	54.8	426	2 I36948	Ig epsilon-chain - chimpanzee (fragment)
2	71	52.6	428	1 EHHU	Ig epsilon chain C
3	66	48.9	107	2 I68730	IgE chain C3 regio
4	66	48.9	107	2 I68726	IgE chain C3 regio
5	66	48.9	388	1 EHW5	Ig epsilon chain C
6	66	48.9	429	1 EHR7	Ig epsilon chain C
7	66	48.9	548	2 S38864	Ig epsilon chain C
8	60	44.4	110	2 S43147	Ig epsilon chain C
9	57	42.2	572	2 B46529	Ig y heavy chain (
10	54	40.0	137	2 S70214	Ig epsilon chain C
11	54	40.0	423	1 EHW55	Ig epsilon chain C
12	54	40.0	598	2 I64042	GTP-binding membra
13	54	40.0	599	1 BVECLA	GTP-binding membra
14	54	40.0	599	2 C91058	GTP-binding elonga
15	54	40.0	599	2 A85903	GTP-binding membra
16	54	40.0	599	2 AE0829	GTP-binding protei
17	54	40.0	602	2 B71561	probable GTPase -
18	54	40.0	602	2 B81714	GTP-binding protei
19	53	39.3	684	2 S60266	novel antigen rece
20	52.5	38.9	450	1 MHOG	Ig mu chain C regi
21	52	38.5	394	2 T19571	hypothetical prote
22	52	38.5	948	2 F87377	GTP-binding protei
23	51	37.8	93	2 I54421	MHC RT1.B-beta2 -
24	51	37.8	220	2 C22360	Ig alpha-2 chain C
25	51	37.8	340	1 A2HU	Ig alpha-2 chain C
26	51	37.8	340	2 B22360	Ig alpha-2 chain C
27	51	37.8	340	2 I56230	Ig mu heavy chain
28	51	37.8	391	1 MHHUBT	Ig mu chain C regi
29	51	37.8	452	1 MHHU	Ig mu chain C regi

30	51	37.8	453	2 S37768	Ig mu chain C regi
31	51	37.8	457	2 S03961	Ig mu chain C regi
32	51	37.8	458	1 MHRB	Ig mu chain C regi
33	51	37.8	473	1 MHHUM	Ig heavy chain - h
34	51	37.8	474	2 S15590	Ig heavy chain - h
35	51	37.8	479	1 MHRBM	Ig mu chain C regi
36	51	37.8	597	2 D82073	GTP-binding protei
37	51	37.8	599	2 AD0331	probable GTP-bindl
38	51	37.8	627	2 S14683	Ig mu chain precu
39	50.5	37.4	893	2 T15183	hypothetical prote
40	50	37.0	328	2 I47161	Ig gamma 3 chain c
41	50	37.0	328	2 I47158	Ig gamma 1 chain c
42	50	37.0	599	2 G83550	GTP-binding protei
43	50	37.0	600	2 H70448	G-protein LepA - A
44	50	37.0	607	2 D86763	GTP-binding protei
45	50	37.0	607	2 B95139	GTP-binding protei

ALIGNMENTS

RESULT 1
I36948
Ig epsilon-chain - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C:Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orang
A:Reference number: I36948; MUID:87147196
A:Accession: I36948
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-426 <RES>
A:Cross-references: GB:M15398; NID:g176797; PIDN:AAA35416.1; PID:g176798
C:Genetics:
A:Introns: 103/1; 209/1; 317/1
C:Superfamily: Immunoglobulin C region; immunoglobulin homology
F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 54.8%; Score 74; DB 2; Length 426;
Best Local Similarity 60.9%; Pred. No. 0.0037;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVRSIAK 24

||||: ||||| ||| :||| |
Db 292 GETYQCRVTHPLPRALVRSTTK 314

RESULT 2

EHHU

Ig epsilon chain C region - human

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1981 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999

C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116;

R:Flanagan, J.G.; Rabbitts, T.H.

EMBO J. 1, 655-660, 1982

A:Title: The sequence of a human immunoglobulin epsilon heavy chain constant region g

A:Reference number: A22771; MUID:84236029

A:Accession: A22771

A:Molecule type: DNA

A:Residues: 1-428 <FLA>

A:Cross-references: GB:I00022; GB:J00227; GB:V00555; NID:g185035

R:Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.

EMBO J. 1, 1539-1544, 1982

A:Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseu

A:Reference number: A23195; MUID:84207910

A:Accession: A23195

A:Molecule type: DNA

A:Residues: 2-428 <UFD>

A:Cross-references: GB:J00222; NID:g184755

R:Zhang, K.; Saxon, A.; Max, E.E.
J. Exp. Med. 176, 233-243, 1992
A:Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing
A:Reference number: PH1214; MUID:92308839
A:Accession: PH1214
A:Molecule type: DNA
A:Residues: 320-428 <ZHA>
A:Cross-references: EMBL:X63693; GB:S38668; NID:g32987
R:Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Kikuchi, M.; Sugii
Nucleic Acids Res. 11, 719-726, 1983
A:Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon chain
A:Reference number: A93491; MUID:83168897
A:Accession: A93491
A:Molecule type: mRNA
A:Residues: 1-428 <SEN>
A:Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035
R:Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.
Cell 29, 691-699, 1982
A:Title: Duplication and deletion in the human immunoglobulin epsilon genes.
A:Reference number: A90824; MUID:83001945
A:Accession: A90824
A:Molecule type: DNA
A:Residues: 1-358, 'L', 360-428 <MAX>
A:Cross-references: GB:J00222; NID:g184755
A:Note: this sequence difference may be due to polymorphism
R:Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
In Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.1-3
A:Reference number: A94418
A:Accession: A94418
A:Molecule type: Protein
A:Residues: 'GAWTL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 125-129, 'L', 131-135, 'L', 137-141, 'L', 143-147, 'L', 149-153, 'L', 155-159, 'L', 161-165, 'L', 167-171, 'L', 173-177, 'L', 179-183, 'L', 185-189, 'L', 191-195, 'L', 197-201, 'L', 203-207, 'L', 209-213, 'L', 215-219, 'L', 221-225, 'L', 227-231, 'L', 233-237, 'L', 239-243, 'L', 245-249, 'L', 251-255, 'L', 257-261, 'L', 263-267, 'L', 269-273, 'L', 275-279, 'L', 281-285, 'L', 287-291, 'L', 293-297, 'L', 299-303, 'L', 305-309, 'L', 311-315, 'L', 317-321, 'L', 323-327, 'L', 329-333, 'L', 335-339, 'L', 341-345, 'L', 347-351, 'L', 353-357, 'L', 359-363, 'L', 365-369, 'L', 371-375, 'L', 377-381, 'L', 383-387, 'L', 389-393, 'L', 395-399, 'L', 401-405, 'L', 407-411, 'L', 413-417, 'L', 419-423, 'L', 425-429, 'L', 431-435, 'L', 437-441, 'L', 443-447, 'L', 449-453, 'L', 455-459, 'L', 461-465, 'L', 467-471, 'L', 473-477, 'L', 479-483, 'L', 485-489, 'L', 491-495, 'L', 497-501, 'L', 503-507, 'L', 509-513, 'L', 515-519, 'L', 521-525, 'L', 527-531, 'L', 533-537, 'L', 539-543, 'L', 545-549, 'L', 551-555, 'L', 557-561, 'L', 563-567, 'L', 569-573, 'L', 575-579, 'L', 581-585, 'L', 587-591, 'L', 593-597, 'L', 599-603, 'L', 605-609, 'L', 611-615, 'L', 617-621, 'L', 623-627, 'L', 629-633, 'L', 635-639, 'L', 641-645, 'L', 647-651, 'L', 653-657, 'L', 659-663, 'L', 665-669, 'L', 671-675, 'L', 677-681, 'L', 683-687, 'L', 689-693, 'L', 695-699, 'L', 701-705, 'L', 707-711, 'L', 713-717, 'L', 719-723, 'L', 725-729, 'L', 731-735, 'L', 737-741, 'L', 743-747, 'L', 749-753, 'L', 755-759, 'L', 761-765, 'L', 767-771, 'L', 773-777, 'L', 779-783, 'L', 785-789, 'L', 791-795, 'L', 797-801, 'L', 803-807, 'L', 809-813, 'L', 815-819, 'L', 821-825, 'L', 827-831, 'L', 833-837, 'L', 839-843, 'L', 845-849, 'L', 851-855, 'L', 857-861, 'L', 863-867, 'L', 869-873, 'L', 875-879, 'L', 881-885, 'L', 887-891, 'L', 893-897, 'L', 899-903, 'L', 905-909, 'L', 911-915, 'L', 917-921, 'L', 923-927, 'L', 929-933, 'L', 935-939, 'L', 941-945, 'L', 947-951, 'L', 953-957, 'L', 959-963, 'L', 965-969, 'L', 971-975, 'L', 977-981, 'L', 983-987, 'L', 989-993, 'L', 995-999, 'L', 1001-1005, 'L', 1007-1011, 'L', 1013-1017, 'L', 1019-1023, 'L', 1025-1029, 'L', 1031-1035, 'L', 1037-1041, 'L', 1043-1047, 'L', 1049-1053, 'L', 1055-1059, 'L', 1061-1065, 'L', 1067-1071, 'L', 1073-1077, 'L', 1079-1083, 'L', 1085-1089, 'L', 1091-1095, 'L', 1097-1101, 'L', 1103-1107, 'L', 1109-1113, 'L', 1115-1119, 'L', 1121-1125, 'L', 1127-1131, 'L', 1133-1137, 'L', 1139-1143, 'L', 1145-1149, 'L', 1151-1155, 'L', 1157-1161, 'L', 1163-1167, 'L', 1169-1173, 'L', 1175-1179, 'L', 1181-1185, 'L', 1187-1191, 'L', 1193-1197, 'L', 1199-1203, 'L', 1205-1209, 'L', 1211-1215, 'L', 1217-1221, 'L', 1223-1227, 'L', 1229-1233, 'L', 1235-1239, 'L', 1241-1245, 'L', 1247-1251, 'L', 1253-1257, 'L', 1259-1263, 'L', 1265-1269, 'L', 1271-1275, 'L', 1277-1281, 'L', 1283-1287, 'L', 1289-1293, 'L', 1295-1299, 'L', 1301-1305, 'L', 1307-1311, 'L', 1313-1317, 'L', 1319-1323, 'L', 1325-1329, 'L', 1331-1335, 'L', 1337-1341, 'L', 1343-1347, 'L', 1349-1353, 'L', 1355-1359, 'L', 1361-1365, 'L', 1367-1371, 'L', 1373-1377, 'L', 1379-1383, 'L', 1385-1389, 'L', 1391-1395, 'L', 1397-1401, 'L', 1403-1407, 'L', 1409-1413, 'L', 1415-1419, 'L', 1421-1425, 'L', 1427-1431, 'L', 1433-1437, 'L', 1439-1443, 'L', 1445-1449, 'L', 1451-1455, 'L', 1457-1461, 'L', 1463-1467, 'L', 1469-1473, 'L', 1475-1479, 'L', 1481-1485, 'L', 1487-1491, 'L', 1493-1497, 'L', 1499-1503, 'L', 1505-1509, 'L', 1511-1515, 'L', 1517-1521, 'L', 1523-1527, 'L', 1529-1533, 'L', 1535-1539, 'L', 1541-1545, 'L', 1547-1551, 'L', 1553-1557, 'L', 1559-1563, 'L', 1565-1569, 'L', 1571-1575, 'L', 1577-1581, 'L', 1583-1587, 'L', 1589-1593, 'L', 1595-1599, 'L', 1601-1605, 'L', 1607-1611, 'L', 1613-1617, 'L', 1619-1623, 'L', 1625-1629, 'L', 1631-1635, 'L', 1637-1641, 'L', 1643-1647, 'L', 1649-1653, 'L', 1655-1659, 'L', 1661-1665, 'L', 1667-1671, 'L', 1673-1677, 'L', 1679-1683, 'L', 1685-1689, 'L', 1691-1695, 'L', 1697-1701, 'L', 1703-1707, 'L', 1709-1713, 'L', 1715-1719, 'L', 1721-1725, 'L', 1727-1731, 'L', 1733-1737, 'L', 1739-1743, 'L', 1745-1749, 'L', 1751-1755, 'L', 1757-1761, 'L', 1763-1767, 'L', 1769-1773, 'L', 1775-1779, 'L', 1781-1785, 'L', 1787-1791, 'L', 1793-1797, 'L', 1799-1803, 'L', 1805-1809, 'L', 1811-1815, 'L', 1817-1821, 'L', 1823-1827, 'L', 1829-1833, 'L', 1835-1839, 'L', 1841-1845, 'L', 1847-1851, 'L', 1853-1857, 'L', 1859-1863, 'L', 1865-1869, 'L', 1871-1875, 'L', 1877-1881, 'L', 1883-1887, 'L', 1889-1893, 'L', 1895-1899, 'L', 1901-1905, 'L', 1907-1911, 'L', 1913-1917, 'L', 1919-1923, 'L', 1925-1929, 'L', 1931-1935, 'L', 1937-1941, 'L', 1943-1947, 'L', 1949-1953, 'L', 1955-1959, 'L', 1961-1965, 'L', 1967-1971, 'L', 1973-1977, 'L', 1979-1983, 'L', 1985-1989, 'L', 1991-1995, 'L', 1997-2001, 'L', 2003-2007, 'L', 2009-2013, 'L', 2015-2019, 'L', 2021-2025, 'L', 2027-2031, 'L', 2033-2037, 'L', 2039-2043, 'L', 2045-2049, 'L', 2051-2055, 'L', 2057-2061, 'L', 2063-2067, 'L', 2069-2073, 'L', 2075-2079, 'L', 2081-2085, 'L', 2087-2091, 'L', 2093-2097, 'L', 2099-2103, 'L', 2105-2109, 'L', 2111-2115, 'L', 2117-2121, 'L', 2123-2127, 'L', 2129-2133, 'L', 2135-2139, 'L', 2141-2145, 'L', 2147-2151, 'L', 2153-2157, 'L', 2159-2163, 'L', 2165-2169, 'L', 2171-2175, 'L', 2177-2181, 'L', 2183-2187, 'L', 2189-2193, 'L', 2195-2199, 'L', 2201-2205, 'L', 2207-2211, 'L', 2213-2217, 'L', 2219-2223, 'L', 2225-2229, 'L', 2231-2235, 'L', 2237-2241, 'L', 2243-2247, 'L', 2249-2253, 'L', 2255-2259, 'L', 2261-2265, 'L', 2267-2271, 'L', 2273-2277, 'L', 2279-2283, 'L', 2285-2289, 'L', 2291-2295, 'L', 2297-2301, 'L', 2303-2307, 'L', 2309-2313, 'L', 2315-2319, 'L', 2321-2325, 'L', 2327-2331, 'L', 2333-2337, 'L', 2339-2343, 'L', 2345-2349, 'L', 2351-2355, 'L', 2357-2361, 'L', 2363-2367, 'L', 2369-2373, 'L', 2375-2379, 'L', 2381-2385, 'L', 2387-2391, 'L', 2393-2397, 'L', 2399-2403, 'L', 2405-2409, 'L', 2411-2415, 'L', 2417-2421, 'L', 2423-2427, 'L', 2429-2433, 'L', 2435-2439, 'L', 2441-2445, 'L', 2447-2451, 'L', 2453-2457, 'L', 2459-2463, 'L', 2465-2469, 'L', 2471-2475, 'L', 2477-2481, 'L', 2483-2487, 'L', 2489-2493, 'L', 2495-2499, 'L', 2501-2505, 'L', 2507-2511, 'L', 2513-2517, 'L', 2519-2523, 'L', 2525-2529, 'L', 2531-2535, 'L', 2537-2541, 'L', 2543-2547, 'L', 2549-2553, 'L', 2555-2559, 'L', 2561-2565, 'L', 2567-2571, 'L', 2573-2577, 'L', 2579-2583, 'L', 2585-2589, 'L', 2591-2595, 'L', 2597-2601, 'L', 2603-2607, 'L', 2609-2613, 'L', 2615-2619, 'L', 2621-2625, 'L', 2627-2631, 'L', 2633-2637, 'L', 2639-2643, 'L', 2645-2649, 'L', 2651-2655, 'L', 2657-2661, 'L', 2663-2667, 'L', 2669-2673, 'L', 2675-2679, 'L', 2681-2685, 'L', 2687-2691, 'L', 2693-2697, 'L', 2699-2703, 'L', 2705-2709, 'L', 2711-2715, 'L', 2717-2721, 'L', 2723-2727, 'L', 2729-2733, 'L', 2735-2739, 'L', 2741-2745, 'L', 2747-2751, 'L', 2753-2757, 'L', 2759-2763, 'L', 2765-2769, 'L', 2771-2775, 'L', 2777-2781, 'L', 2783-2787, 'L', 2789-2793, 'L', 2795-2799, 'L', 2801-2805, 'L', 2807-2811, 'L', 2813-2817, 'L', 2819-2823, 'L', 2825-2829, 'L', 2831-2835, 'L', 2837-2841, 'L', 2843-2847, 'L', 2849-2853, 'L', 2855-2859, 'L', 2861-2865, 'L', 2867-2871, 'L', 2873-2877, 'L', 2879-2883, 'L', 2885-2889, 'L', 2891-2895, 'L', 2897-2901, 'L', 2903-2907, 'L', 2909-2913, 'L', 2915-2919, 'L', 2921-2925, 'L', 2927-2931, 'L', 2933-2937, 'L', 2939-2943, 'L', 2945-2949, 'L', 2951-2955, 'L', 2957-2961, 'L', 2963-2967, 'L', 2969-2973, 'L', 2975-2979, 'L', 2981-2985, 'L', 2987-2991, 'L', 2993-2997, 'L', 2999-3003, 'L', 3005-3009, 'L', 3011-3015, 'L', 3017-3021, 'L', 3023-3027, 'L', 3029-3033, 'L', 3035-3039, 'L', 3041-3045, 'L', 3047-3051, 'L', 3053-3057, 'L', 3059-3063, 'L', 3065-3069, 'L', 3071-3075, 'L', 3077-3081, 'L', 3083-3087, 'L', 3089-3093, 'L', 3095-3099, 'L', 3101-3105, 'L', 3107-3111, 'L', 3113-3117, 'L', 3119-3123, 'L', 3125-3129, 'L', 3131-3135, 'L', 3137-3141, 'L', 3143-3147, 'L', 3149-3153, 'L', 3155-3159, 'L', 3161-3165, 'L', 3167-3171, 'L', 3173-3177, 'L', 3179-3183, 'L', 3185-3189, 'L', 3191-3195, 'L', 3197-3201, 'L', 3203-3207, 'L', 3209-3213, 'L', 3215-3219, 'L', 3221-3225, 'L', 3227-3231, 'L', 3233-3237, 'L', 3239-3243, 'L', 3245-3249, 'L', 3251-3255, 'L', 3257-3261, 'L', 3263-3267, 'L', 3269-3273, 'L', 3275-3279, 'L', 3281-3285, 'L', 3287-3291, 'L', 3293-3297, 'L', 3299-3303, 'L', 3305-3309, 'L', 3311-3315, 'L', 3317-3321, 'L', 3323-3327, 'L', 3329-3333, 'L', 3335-3339, 'L', 3341-3345, 'L', 3347-3351, 'L', 3353-3357, 'L', 3359-3363, 'L', 3365-3369, 'L', 3371-3375, 'L', 3377-3381, 'L', 3383-3387, 'L', 3389-3393, 'L', 3395-3399, 'L', 3401-3405, 'L', 3407-3411, 'L', 3413-3417, 'L', 3419-3423, 'L', 3425-3429, 'L', 3431-3435, 'L', 3437-3441, 'L', 3443-3447, 'L', 3449-3453, 'L', 3455-3459, 'L', 3461-3465, 'L', 3467-3471, 'L', 3473-3477, 'L', 3479-3483, 'L', 3485-3489, 'L', 3491-3495, 'L', 3497-3501, 'L', 3503-3507, 'L', 3509-3513, 'L', 3515-3519, 'L', 3521-3525, 'L', 3527-3531, 'L', 3533-3537, 'L', 3539-3543, 'L', 3545-3549, 'L', 3551-3555, 'L', 3557-3561, 'L', 3563-3567, 'L', 3569-3573, 'L', 3575-3579, 'L', 3581-3585, 'L', 3587-3591, 'L', 3593-3597, 'L', 3599-3603, 'L', 3605-3609, 'L', 3611-3615, 'L', 3617-3621, 'L', 3623-3627, 'L', 3629-3633, 'L', 3635-3639, 'L', 3641-3645, 'L', 3647-3651, 'L', 3653-3657, 'L', 3659-3663, 'L', 3665-3669, 'L', 3671-3675, 'L', 3677-3681, 'L', 3683-3687, 'L', 3689-3693, 'L', 3695-3699, 'L', 3701-3705, 'L', 3707-3711, 'L', 3713-3717, 'L', 3719-3723, 'L', 3725-3729, 'L', 3731-3735, 'L', 3737-3741, 'L', 3743-3747, 'L', 3749-3753, 'L', 3755-3759, 'L', 3761-3765, 'L', 3767-3771, 'L', 3773-3777, 'L', 3779-3783, 'L', 3785-3789, 'L', 3791-3795, 'L', 3797-3801, 'L', 3803-3807, 'L', 3809-3813, 'L', 3815-3819, 'L', 3821-3825, 'L', 3827-3831, 'L', 3833-3837, 'L', 3839-3843, 'L', 3845-3849, 'L', 3851-3855, 'L', 3857-3861, 'L', 3863-3867, 'L', 3869-3873, 'L', 3875-3879, 'L', 3881-3885, 'L', 3887-3891, 'L', 3893-3897, 'L', 3899-3903, 'L', 3905-3909, 'L', 3911-3915, 'L', 3917-3921, 'L', 3923-3927, 'L', 3929-3933, 'L', 3935-3939, 'L', 3941-3945, 'L', 3947-3951, 'L', 3953-3957, 'L', 3959-3963, 'L', 3965-3969, 'L', 3971-3975, 'L', 3977-3981, 'L', 3983-3987, 'L', 3989-3993, 'L', 3995-3999, 'L', 4001-4005, 'L', 4007-4011, 'L', 4013-4017, 'L', 4019-4023, 'L', 4025-4029, 'L', 4031-4035, 'L', 4037-4041, 'L', 4043-4047, 'L', 4049-4053, 'L', 4055-4059, 'L', 4061-4065, 'L', 4067-4071, 'L', 4073-4077, 'L', 4079-4083, 'L', 4085-4089, 'L', 4091-4095, 'L', 4097-4101, 'L', 4103-4107, 'L', 4109-4113, 'L', 4115-4119, 'L', 4121-4125, 'L', 4127-4131, 'L', 4133-4137, 'L', 4139-4143, 'L', 4145-4149, 'L', 4151-4155, 'L', 4157-4161, 'L', 4163-4167, 'L', 4169-4173, 'L', 4175-4179, 'L', 4181-4185, 'L', 4187-4191, 'L', 4193-4197, 'L', 4199-4203, 'L', 4205-4209, 'L', 4211-4215, 'L', 4217-4221, 'L', 4223-4227, 'L', 4229-4233, 'L', 4235-4239, 'L', 4241-4245, 'L', 4247-4251, 'L', 4253-4257, 'L', 4259-4263, 'L', 4265-4269, 'L', 4271-4275, 'L', 4277-4281, 'L', 4283-4287, 'L', 4289-4293, 'L', 4295-4299, 'L', 4301-4305, 'L', 4307-4311, 'L', 4313-4317, 'L', 4319-4323, 'L', 4325-4329, 'L', 4331-4335, 'L', 4337-4341, 'L', 4343-4347, 'L', 4349-4353, 'L', 4355-4359, 'L', 4361-4365, 'L', 4367-4371, 'L', 4373-4377, 'L', 4379-4383, 'L', 4385-4389, 'L', 4391-4395, 'L', 4397-4401, 'L', 4403-4407, 'L', 4409-4413, 'L', 4415-4419, 'L', 4421-4425, 'L', 4427-4431, 'L', 4433-4437, 'L', 4439-4443, 'L', 4445-4449, 'L', 4451-4455, 'L', 4457-4461, 'L', 4463-4467, 'L', 4469-4473, 'L', 4475-4479, 'L', 4481-4485, 'L', 4487-4491, 'L', 4493-4497, 'L', 4499-4503, 'L', 4505-4509, 'L', 4511-4515, 'L', 4517-4521, 'L', 4523-4527, 'L', 4529-4533, 'L', 4535-4539, 'L', 4541-4545, 'L', 4547-4551, 'L', 4553-4557, 'L', 4559-4563, 'L', 4565-4569, 'L', 4571-4575, 'L', 4577-4581, 'L', 4583-4587, 'L', 4589-4593, 'L', 4595-4599, 'L', 4601-4605, 'L', 4607-4611, 'L', 4613-4617, 'L', 4619-4623, 'L', 4625-4629, 'L', 4631-4635, 'L', 4637-4641, 'L', 4643-4647, 'L', 4649-4653, 'L', 4655-4659, 'L', 4661-4665, 'L', 4667-4671, 'L', 4673-4677, 'L', 4679-4683, 'L', 4685-4689, 'L', 4691-4695, 'L', 4697-4701, 'L', 4703-4707, 'L', 4709-4713, 'L', 4715-4719, 'L', 4721-4725, 'L', 4727-4731, 'L', 4733-4737, 'L', 4739-4743, 'L', 4745-4749, 'L', 4751-4755, 'L', 4757-4761, 'L', 4763-4767, 'L', 4769-4773, 'L', 4775-4779, 'L', 4781-4785, 'L', 4787-4791, 'L', 4793-4797, 'L', 4799-4803, 'L', 4805-4809, 'L', 4811-4815, 'L', 4817-4821, 'L', 4823-4827, 'L', 4829-4833, 'L', 4835-4839, 'L', 4841-4845, 'L', 4847-4851, 'L', 4853-4857, 'L', 4859-4863, 'L', 4865-4869, 'L', 4871-4875, 'L', 4877-4881, 'L', 4883-4887, 'L', 4889-4893, 'L', 4895-4899, 'L', 4901-4905, 'L', 4907-4911, 'L', 4913-4917, 'L', 4919-4923, 'L', 4925-4929, 'L', 4931-4935, 'L', 4937-4941, 'L', 4943-4947, 'L', 4949-4953, 'L', 4955-4959, 'L', 4961-4965, 'L', 4967-4971, 'L', 4973-4977, 'L', 4979-4983, 'L', 4985-4989, 'L', 4991-4995, 'L', 4997-5001, 'L', 5003-5007, 'L', 5009-5013, 'L', 5015-5019, 'L', 5021-5025, 'L', 5027-5031, 'L', 5033-5037, 'L', 5039-5043, 'L', 5045-5049, 'L', 5051-5055, 'L', 5057-5061, 'L', 5063-5067, 'L', 5069-5073, 'L', 5075-5079, 'L', 5081-5085, 'L', 5087-5091, 'L', 5093-5097, 'L', 5099-5103, 'L', 5105-5109, 'L', 5111-5115, 'L', 5117-5121, 'L', 5123-5127, 'L', 5129-5133, 'L', 5135-5139, 'L', 5141-5145, 'L', 5147-5151, 'L', 5153-5157, 'L', 5159-5163, 'L', 5165-5169, 'L', 5171-5175, 'L', 5177-5181, 'L', 5183-5

A;Contents: myeloma IR162							
A;Accession: A90937							
A:Molecule type: mRNA							
C;Residues: 'N',169-307,'L',309-342 <KIN>							
C;Complex: An immunoglobulin heterotrimer subunit consists of two identical light chains disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C;Superfamily: immunoglobulin C region; immunoglobulin homology							
C;Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin							
F;19-80/Domain: immunoglobulin homology <IM1>							
F;118-186/Domain: immunoglobulin homology <IM2>							
F;223-291/Domain: immunoglobulin homology <IM3>							
F;327-398/Domain: immunoglobulin homology <IM4>							
F;46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predicted							
Query Match	48.9%;	Score 66;	DB 1;	Length 429;			
Best Local Similarity	52.2%;	Pred. No. 0.058;					
Matches 12;	Conservative 3;	Mismatches 8;	Indels 0;	Gaps 0;			
QY 2 GETYSTVSHPDLPREVVRSTAK 24							
Dd 284 GEGYQCRVDHPHPPKPIVRSITK 306							
RESULT 7							
S38864							
Ig epsilon chain C region - mouse (fragment)							
C;Species: Mus musculus (house mouse)							
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001							
C;Accession: S38864							
R;Kipp, B.; Becker, W.; Schlaack, M.							
submitted to the EMBL Data Library, November 1993							
A;Description: Combination of a defined specificity and desired isotype by cloning of							
A;Reference number: S38864							
A;Accession: S38864							
A>Status: preliminary							
A:Molecule type: mRNA							
A;Residues: 1-548 <KIP>							
A;Cross-references: EMBL:X77397; NID:g416537; PIDN:CAA81788.1; PID:g940782							
C;Superfamily: immunoglobulin C region; immunoglobulin homology							
F;353-421/Domain: immunoglobulin homology <IMM>							
Query Match	48.9%;	Score 66;	DB 2;	Length 548;			
Best Local Similarity	52.2%;	Pred. No. 0.076;					
Matches 12;	Conservative 3;	Mismatches 8;	Indels 0;	Gaps 0;			
QY 2 GETYSTVSHPDLPREVVRSTAK 24							
Dd 414 GYGQCIVDHPDPKPPIVRSITK 436							
RESULT 8							
S43147							
Ig upsilon chain - duck (fragment)							
C;Species: Anas platyrhynchos (domestic duck)							
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 11-Jan-2000							
C;Accession: S43147							
R;Magor, K.E.; Higgins, D.A.; Middleton, D.L.; Warr, G.W.							
submitted to the EMBL Data Library, March 1994							
A;Description: Evidence from duck immunoglobulin genes that IgY is the common ancestor							
A;Reference number: S43145							
A;Accession: S43147							
A>Status: preliminary							
A:Molecule type: DNA							
A;Residues: 1-110 <MAG>							
A;Cross-references: EMBL:X78355; NID:g468612; PID:g468613							
C;Superfamily: immunoglobulin C region; immunoglobulin homology							
Query Match	44.4%;	Score 60;	DB 2;	Length 110;			
Best Local Similarity	52.2%;	Pred. No. 0.11;					
Matches 12;	Conservative 3;	Mismatches 8;	Indels 0;	Gaps 0;			

QY 2 GETYKSTVSHPDLPREVVRSAK 24
 || : || | ||| : |||||

Db 86 GERFTCTVQHEDLPVPLGKSAK 108

RESULT 9

B46529
 Ig Y heavy chain (7.8S) - duck
 N Alternate names: Ig gamma chain (7.8S)
 C:Species: Anas platyrhynchos (domestic duck)
 C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 C:Accession: B46529; S20759
 R:Magor, K.E.; Warr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.
 J. Immunol. 149, 2627-2633, 1992
 A:Title: Structural relationship between the two IgY of the duck, Anas platyrhynchos: m
 A:Reference number: A46529; MUID:93017865
 A:Accession: B46529
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-572 <MAG>
 A:Cross-references: EMBL:X65219; NID:g62442; PIDN:CAA46322.1; PID:g62443
 A:Experimental source: spleen
 A>Note: sequence extracted from NCBI backbone (NCBIP:116127)
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:37-120/Domain: immunoglobulin homology <IMM>

Query Match 42.2%; Score 57; DB 2; Length 572;

Best Local Similarity 52.2%; Pred. No. 1.8;

Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVVRSAK 24

|| : || | ||| : |||||

Db 437 GERFTCTVQHEDLPVPLGKSAK 459

RESULT 10

S70214
 lepA protein homolog - Photobacterium sp. (fragment)
 C:Species: Photobacterium sp.
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 02-Feb-2001
 C:Accession: S70214
 R:Chi, E.; Bartlett, D.H.
 Mol. Microbiol. 17, 713-726, 1995
 A:Title: An rpoE-like locus controls outer membrane protein synthesis and growth at cold
 A:Reference number: S70210; MUID:96111491
 A:Accession: S70214
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-137 <CHT>

A:Cross-references: EMBL:L41667; NID:g77746; PIDN:AAB31927.1; PID:g777751
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
 C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homol
 C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
 F:5-134/Domain: translation elongation factor Tu homology <ETU>
 F:11-18/Region: nucleotide-binding motif A (P-loop)
 F:131-134/Region: GTP-binding NKXD motif

Query Match 40.0%; Score 54; DB 2; Length 137;

Best Local Similarity 46.4%; Pred. No. 1.1;

Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK-----STVSHPDLPREVVRSAK 25

|||| : || | ||| : |||||

Db 68 GETYQLNFIDTPGHVDFSVESRSLAAC 95

RESULT 11

EHMS
 Ig epsilon chain C region (version 2) - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
 C:Accession: A02145
 R:Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.
 EMBO J. 1, 1117-1123, 1982
 A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison
 A:Reference number: A90966; MUID:84236092
 A:Accession: A02145
 A:Molecule type: DNA
 A:Residues: 1-423 <ISH>

A>Note: the sequence was determined from the germline gene

C:Genetics: 91

A:Introns: 91/1; 199/1; 307/1

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:16-77/Domain: immunoglobulin homology <IMM1>
 F:115-183/Domain: immunoglobulin homology <IMM2>
 F:220-288/Domain: immunoglobulin homology <IMM3>
 F:325-396/Domain: immunoglobulin homology <IMM4>
 F:23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
 F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 54; DB 1; Length 423;

Best Local Similarity 47.6%; Pred. No. 3.6;

Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVVRSI 22

| : | | | : ||||

Db 281 GYGQCVDPRDPFKPIVRSI 301

RESULT 12

I84042
 GTP-binding membrane protein lepA - Haemophilus influenzae (straln Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
 C:Accession: I64042
 R:Flaischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630

A:Accession: I64042

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-598 <TIGR>

A:Cross-references: GB:U32687; GB:L42023; NID:g1572955; PIDN:AAC21694.1; PID:g1572960

C:Genetics:

A:Gene: lepA

C:Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu ho

C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop

F:5-134/Domain: translation elongation factor Tu homology <ETU>

F:11-18/Region: nucleotide-binding motif A (P-loop)

F:131-134/Region: GTP-binding NKXD motif

F:162-164/Region: GTP-binding SAK/L motif

F:17,18,53,131,132,134,162/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

Query Match 40.0%; Score 54; DB 2; Length 598;

Best Local Similarity 46.4%; Pred. No. 5.2;

Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK-----STVSHPDLPREVVRSAK 25

|||| : || | ||| : |||||

Db 68 GETYQLNFIDTPGHVDFSVESRSLAAC 95

RESULT 13

```

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC LEPA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006042; AAK02147.1; -.
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR Pfam: PF00009; GTP_EFTU.1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR PROSITE: PS00301; EFACOR_GTP; 1.
KW GTP-binding; Complete proteome.
FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
SQ SEQUENCE 598 AA; 66211 MW; 4D41723C2AEBFD10 CRC64;

Query Match 40.0%; Score 54; DB 1; Length 598;
Best Local Similarity 46.4%; Pred. No. 1.5;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

OY 2 GETYK----STVSHPDLPREVVRSTAK 25
||||| | | | | | | | | |
Db 68 GETYQLNFIDTPCHVDVFSYVSKSLAAC 95

RESULT 6
LEPA_ECOLI STANDARD; PRT; 599 AA.
AC P07682; P76590;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR B2569 OR Z3851 OR ECS3435.
OS Escherichia coli, and
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85207751; PubMed=2987248;
RA March P.E., Inouye M.;
RT "Characterization of the lep operon of Escherichia coli.
RT Identification of the promoter and the gene upstream of the signal
RT peptidase I gene."
RL J. Biol. Chem. 260:7206-7213(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mañ B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN [6]
RP SIMILARITY TO GTP-BINDING PROTEINS.
RX MEDLINE=86067932; PubMed=2999765;
RA March P.E., Inouye M.;
RT "GTP-binding membrane protein of Escherichia coli with sequence
RT homology to initiation factor 2 and elongation factors Tu and G."
RL Proc. Natl. Acad. Sci. U.S.A. 82:7500-7504(1985).
RN [7]
RP SIMILARITY TO GTP-BINDING PROTEINS.
RX MEDLINE=88082865; PubMed=2826164;
RA Kushiro A., Shimizu M., Tomita K.;
RT "Molecular cloning and sequence determination of the tuf gene coding
RT for the elongation factor Tu of Thermus thermophilus HB8."
RL Eur. J. Biochem. 170:93-98(1987).
CC -!- SUBCELLULAR LOCATION: PRESENT IN THE CYTOPLASMIC MEMBRANE AND IS
CC ALSO FOUND IN THE PERIPLASM.
CC -!- MISCELLANEOUS: THE OVERPRODUCTION OF THIS PROTEIN IS LETHAL TO
CC E. COLI.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC LEPA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K00426; AAA24063.1; -.
DR EMBL: D64044; BAA10916.1; -.
DR EMBL: AE000343; AAC75622.1; -.
DR EMBL: AE005487; AAG57685.1; -.
DR EMBL: AP002562; BAB36858.1; -.
DR PIR: A22627; BVCELA.
DR HSP: P13551; 1DAR.
DR EcoGene: EG10529; lepa.
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PROSITE: PS00301; EFACOR_GTP; 1.
KW Membrane; GTP-binding; Complete proteome.
FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT NP_BIND 131 134 GTP (BY SIMILARITY).
FT CONFLICT 156 185
FT CONFLICT 266 266 MISSING (IN REF. 1 AND 2).
FT CONFLICT 599 AA; 66570 MW; 91E3678557A46D49 CRC64;
SQ SEQUENCE

```



```

Query Match          40.0%; Score 54; DB 1; Length 599;
Best Local Similarity 46.4%; Pred. NO. 1.5;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY      2  GENYK-----STVSHPDLPREVVSIAKC 25
      ||||:  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      68  GETYQLNFIDTPGHVDFSEVSRSLAAC 95

RESULT 7
LEPA_SALTY
ID      LEPA_SALTY      STANDARD;      PRT;      599 AA.
AC      F23698;
DT      01-NOV-1991 (Rel. 20, Created)
DT      01-MAR-2002 (Rel. 41, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      GTP-binding protein lepa.
DE      LEPA OR STM2583 OR STY2829.
GN      Salmonella typhimurium, and
OS      Salmonella typhi.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Salmonella.
OX      NCBI_TaxID=602, 601;
[1]
SEQUENCE FROM N.A.
RC      SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX      MEDLINE=21534948; PubMed=11677609;
RA      McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA      Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA      Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA      Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA      Waterston R., Wilson R.K.;
RA      *Complete genome sequence of Salmonella enterica serovar Typhimurium
RT      LT2";
RL      Nature 413:852-856(2001).
[2]
SEQUENCE OF 565-599 FROM N.A.
RC      SPECIES=S.typhimurium; STRAIN=JM83;
RX      MEDLINE=91066835; PubMed=2250650;
RA      van Dijk J.M., van den Bergh R., Reversma T., Smith H., Bron S.,
RA      Venema G.;
RT      "Molecular cloning of the Salmonella typhimurium lep gene in
RT      Escherichia coli.";
RL      Mol. Gen. Genet. 223:233-240(1990).
[3]
SEQUENCE FROM N.A.
RC      SPECIES=S.typhi; STRAIN=CTL18;
RX      MEDLINE=21534947; PubMed=11677608;
RA      Parkhill J., Dougan G., James K.D., Bentley S.D., Holden M.T.G., Sebahia M.,
RA      Churche C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA      Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA      Cronin A., Davis P., Davies R.M., Bowd L., White N., Farrar J.,
RA      Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA      Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA      Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA      Whitehead S., Barrall B.G.;
RT      *Complete genome sequence of a multiple drug resistant Salmonella
RT      enterica serovar typhi CTL18.";
RL      Nature 413:848-852(2001).
CC      -1- SUBCELLULAR LOCATION: PRESENT IN THE CYTOPLASMIC MEMBRANE AND IS
CC      ALSO FOUND IN THE PERIPLASM.
CC      -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC      LEPA SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
```

```

CC EMBL; AE008817; AAL21477.1; -.
DR DR EMBL; X54933; CAA38695.1; -.
DR DR EMBL; AL627275; CAD02785.1; -.
DR PIR; S12019; S12019.
DR DR StyGene; SG10193; lepa.
DR DR InterPro; IPR000795; GTP_EFTU.
DR DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Membrane; GTP-binding; Complete proteome.
FT NP_BIND 11 18 GTP (BY SIMILARITY).
FT FT NP_BIND 77 81 GTP (BY SIMILARITY).
FT FT NP_BIND 131 134 GTP (BY SIMILARITY).
SQ SEQUENCE 599 AA; 66538 MW; DOBD7FBE947157DA CRC64;

Query Match 40.0%; Score 54; DB 1; Length 599;
Best Local Similarity 46.4%; Pred. No. 1.5;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

Qy 2 GETYK----STVSHPDLPREVRSIAKC 25
      ||||:  |  |  |  ||||:  |
Db 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95

RESULT 8
LEPA_CHLMU
ID LEPA_CHLMU STANDARD; PRT; 602 AA.
AC Q9PKX6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR TC0334.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mopn / Nigg.
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RT pneumoniae AR39."
RT Nucleic Acids Res. 28:1397-1406(2000).
RL CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC LEPA SUBFAMILY.
CC CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; AE002301; AAF39197.1; -.
DR TIGR; TC0334; -.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR003575; Small_GTPase.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR SMART; SM00010; small_GTPase; 1.
DR PROSITE; PS00301; EFACOR_GTP; FALSE_NEG.
DR GTP-binding; Complete proteome.
KW NP_BIND 16 23 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 135 138 GTP (BY SIMILARITY).

```

SQ SEQUENCE 602 AA; 67378 MW; 1EDA355FF60EBF3D CRC64;

Query Match 40.0%; Score 54; DB 1; Length 602;
Best Local Similarity 46.4%; Pred. No. 1.5;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

OY 2 GETYK----STVSHPDLPREVVRSTAKC 25

||||: | | | | | | | | | |

Db 72 GETYELNLIDTPGHVDFSVSRSLAAC 99

RESULT 9

ID LEPA_CHLTR STANDARD; PRT; 602 AA.

AC 084067;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE GTP-binding protein lepa.

GN LEPA OR CT064.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D/UV-3/CX;

RX MEDLINE=99000809; PubMed=9784136;

RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,

RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,

RA Davis R.W.;

RT "Genome sequence of an obligate intracellular pathogen of humans:

RT Chlamydia trachomatis";

RL Science 282:754-759(1998).

CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.

CC LEPA SUBFAMILY.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AE001281; AAC67655.1; -

DR HSSP; P13551; 1ELO.

DR InterPro; IPR000795; GTP_EFTU.

DR InterPro; IPR004161; GTP_EFTU_D2.

DR Pfam; PF00009; GTP_EFTU; 1.

DR Pfam; PF03144; GTP_EFTU_D2; 1.

DR PRINTS; PR00315; ELONGATNFCT.

DR PROSITE; PS00301; EFATOR.GTP; FALSE_NEG.

KW GTP-binding; Complete proteome.

FT NP_BIND 16 23 GTP (BY SIMILARITY).

FT NP_BIND 81 85 GTP (BY SIMILARITY).

FT NP_BIND 135 138 GTP (BY SIMILARITY).

SQ SEQUENCE 602 AA; 67435 MW; 98F67C990196EEC4 CRC64;

Query Match 40.0%; Score 54; DB 1; Length 602;

Best Local Similarity 46.4%; Pred. No. 1.5;

Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

OY 2 GETYK----STVSHPDLPREVVRSTAKC 25

||||: | | | | | | | | | |

Db 72 GETYELNLIDTPGHVDFSVSRSLAAC 99

RESULT '10

MUC_CANFA

ID MUC_CANFA

AC P01874; STANDARD; PRT; 450 AA.

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig mu chain C region.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RP SEQUENCE OF 1-177 (MOO).

RX MEDLINE=80077682; PubMed=117299;

RA McCumber L.J., Capra J.D.;

RT "The complete amino-acid sequence of a canine mu chain.";

RL Mol. Immunol. 16:565-570(1979).

RN [2]

RP SEQUENCE OF 178-450 (MOO).

RX MEDLINE=78180587; PubMed=653360;

RA Wasserman R.L., Capra J.D.;

RT "Amino acid sequence of the Fc region of a canine immunoglobulin M:

RT interspecies homology for the Igm class.";

RL Science 200:1159-1161(1978).

DR PIR; A02169; MHDG.

DR HSSP; P01857; 1FC1.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003597; Ig_cl.

DR InterPro; IPR003600; Ig_like.

DR Pfam; PF00047; Ig; 3.

DR SMART; SM00410; IG_Like; 2.

DR SMART; SM00407; IGC1; 2.

DR PROSITE; PS00290; IG_MHC; 3.

KW Immunoglobulin domain; Immunoglobulin C region.

FT NON_TER 1

SQ SEQUENCE 450 AA; 48895 MW; 9D460DA9D1012F5D CRC64;

Query Match 38.9%; Score 52.5; DB 1; Length 450;

Best Local Similarity 50.0%; Pred. No. 1.9;

Matches 12; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

OY 2 GETYKSTVSHPDLPREVVRSTAK 24

||||: | | | | | | | | | |

Db 405 GETYTCVVAHESLPNRTVSVDK 428

RESULT 11

ALC2_HUMAN

ID ALC2_HUMAN STANDARD; PRT; 340 AA.

AC P01877;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ig alpha-2 chain C region.

GN IGH2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84130179; PubMed=6421489;

RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;

RT "Mechanisms of divergence and convergence of the human immunoglobulin

RT alpha 1 and alpha 2 constant region gene sequences.";

RL Cell 36:681-688(1984).

RN [2]

RP SEQUENCE (BUT).

RX MEDLINE=78137069; PubMed=416441;

RA Torano A., Putnam F.W.;

RT "Complete amino acid sequence of the alpha 2 heavy chain of a human

RT IgA2 immunoglobulin of the A2m (2) allotype.";

RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).

RN [3]

RP SEQUENCE (MYELOMA PROTEIN LAN).


```

FT DISULFID 13 13 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 27 89 BY SIMILARITY.
FT DISULFID 136 200 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 217 217 BY SIMILARITY.
FT DISULFID 248 307 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 295 295 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 355 417 BY SIMILARITY.
FT DISULFID 456 456 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 283 283 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 457 AA; 50074 MW; 56C8C086DA4462E9 CRC64;

Query Match 37.8%; Score 51; DB 1; Length 457;
Best Local Similarity 39.1%; Pred. No. 3.2;
Matches 9; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2 GETYKSTVSHPDLPREVVRISAK 24
|| : || ||| : ||::
Db 302 GEKFTCTVQHSDPLSPKQSLSR 324

```

RESULT 15

```

MUC_RABIT
ID MUC_RABIT STANDARD; PRT; 458 AA.
AC P03988;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 36, Last annotation update)
DE Ig M chain C region.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A. (A2 ALLOTYPED).
RX MEDLINE=84088930; PubMed=6418803;
RA Bernstein K.E., Alexander C.B., Reddy E.P., Mage R.G.;
RT "Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain
RT of Vha2 allotype: comparisons with Vha1 and membrane mu sequences.";
RL J. Immunol. 132:490-495(1984).
CC -!- ALTERNATIVE PRODUCTS: DURING DIFFERENTIATION, B LYMPHOCYTES
CC SWITCH FROM EXPRESSION OF MEMBRANE-BOUND IGM TO SECRETION OF IGM.
CC THE MU CHAINS OF MEMBRANE AND SECRETED IGM DIFFER IN THEIR C-
CC TERMINAL SEGMENTS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K01357; -; NOT_ANNOTATED_CDS.
CC PIR; A02164; MHRB.
CC HSP; P13760; 2SEB.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR003600; Ig_like.
CC Pfam; PF00047; Ig; 4.
CC SMART; SM00410; IG_like; 2.
CC SMART; SM00407; IGcl; 2.
CC PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1 106 CH1.
FT DOMAIN 107 222 CH2.

```

```

FT DOMAIN 223 327 CH3.
FT DISULFID 328 458 CH4.
FT DISULFID 14 14 INTERCHAIN (WITH LIGHT CHAIN) (PROBABLE).
FT DISULFID 28 90 BY SIMILARITY.
FT DISULFID 137 200 BY SIMILARITY.
FT DISULFID 219 219 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 249 308 BY SIMILARITY.
FT DISULFID 296 308 INTERCHAIN (WITH HEAVY CHAIN) (PROBABLE).
FT DISULFID 356 418 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 458 AA; 49897 MW; F338D6A3B28E4864 CRC64;

Query Match 37.8%; Score 51; DB 1; Length 458;
Best Local Similarity 43.5%; Pred. No. 3.2;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 2 GETYKSTVSHPDLPREVVRISAK 24
|| : || ||| : ||:|
Db 303 GEQFTCTVTHADLPPLKHTISK 325

```

Search completed: August 26, 2002, 10:17:04
Job time: 700 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:16:21 ; Search time 66.82 Seconds
(without alignments)
64.724 Million cell updates/sec

Title: US-09-701-623C-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP TREMBL.19.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mammal.*
8: sp_mhc.*
9: sp_organelle.*
10: sp_phage.*
11: sp_plant.*
12: sp_rodent.*
13: sp_virus.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	40.0	137	2 Q57464	Q57464 photobacter
2	53	39.3	684	13 Q90544	Q90544 ginglymosto
3	52	38.5	394	5 Q17620	Q17620 caenorhabdi
4	52	38.5	648	16 Q9A9P4	Q9A9P4 caulobacter
5	51	37.8	93	7 Q31260	Q31260 rattus norv
6	51	37.8	341	11 Q08284	Q08284 mus musculu
7	51	37.8	369	16 Q9CM48	Q9CM48 pasteurella
8	51	37.8	375	4 Q9BS21	Q9BS21 homo sapien
9	51	37.8	416	4 Q9NPP6	Q9NPP6 homo sapien
10	51	37.8	557	11 Q99K47	Q99K47 mus musculu
11	51	37.8	597	4 Q9BQB8	Q9BQB8 homo sapien
12	51	37.8	597	4 Q9BU10	Q9BU10 homo sapien
13	51	37.8	597	4 Q96BB9	Q96BB9 homo sapien
14	51	37.8	597	16 Q9KPB0	Q9KPB0 vibrio chol
15	51	37.8	613	4 Q96YV0	Q96YV0 homo sapien
16	51	37.8	614	4 Q96GA6	Q96GA6 homo sapien

17	51	37.8	618	4 Q96AA6	Q96AA6 homo sapien
18	51	37.8	882	5 Q9NE67	Q9NE67 leishmania
19	50.5	37.4	893	5 Q02099	Q02099 caenorhabdi
20	50	37.0	245	2 Q9EV20	Q9EV20 listeria mo
21	50	37.0	599	16 Q915G8	Q915G8 pseudomonas
22	50	37.0	607	16 Q97OK5	Q97OK5 streptococc
23	50	37.0	608	16 Q92BN4	Q92BN4 listeria in
24	50	37.0	610	16 Q99ZV8	Q99ZV8 streptococc
25	49	36.3	473	11 Q91Z05	Q91Z05 mus musculu
26	49	36.3	520	10 Q48786	Q48786 arabidopsis
27	49	36.3	712	5 Q960E2	Q960E2 drosophila
28	49	36.3	1105	5 Q9VN44	Q9VN44 drosophila
29	48.5	35.9	279	11 Q9D9F8	Q9D9F8 mus musculu
30	48	35.6	335	11 Q08283	Q08283 mesocricetu
31	48	35.6	425	11 Q9DB22	Q9DB22 mus musculu
32	48	35.6	608	16 Q92SU3	Q92SU3 rhizobium m
33	48	35.6	962	5 Q9GYU5	Q9GYU5 trypanosoma
34	48	35.6	1048	13 P79954	P79954 xenopus lae
35	47	34.8	112	16 Q981E5	Q981E5 rhizobium l
36	47	34.8	217	7 Q30827	Q30827 ovis aries
37	47	34.8	245	7 Q31270	Q31270 rattus norv
38	47	34.8	261	7 Q9TQA7	Q9TQA7 rattus norv
39	47	34.8	264	7 Q9TQA5	Q9TQA5 rattus norv
40	47	34.8	364	16 Q97PS4	Q97PS4 streptococc
41	47	34.8	387	5 Q17621	Q17621 caenorhabdi
42	47	34.8	479	10 Q9LWH4	Q9LWH4 oryza sativ
43	47	34.8	607	16 Q99TR4	Q99TR4 staphylococ
44	47	34.8	3134	5 Q25994	Q25994 plasmodium
45	46	34.1	176	2 Q9R8Y8	Q9R8Y8 rhizobium s

ALIGNMENTS

RESULT 1
Q57464
ID Q57464 PRELIMINARY; PRT; 137 AA.
AC Q57464;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RPOE, LEPA AND ORF 2, 3, 4 GENES (FRAGMENT).
GN LEPA.
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
OC Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SS9;
RA Chi E., Bartlett D.H.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SS9;
RA Chi E., Bartlett D.H.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; L41688; AAA65230.1; -.
DR EMBL; L41667; AAB31927.1; -.
DR HSP; P13551; 1FNM.
DR InterPro; IPR000795; GTP_EFTU.
DR Pfam; PF00009; GTP_EFTU; 1.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW GTP-binding; Protein biosynthesis.
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 15016 MW; 7FC5E05FD6AF1F3E CRC64;

Query Match 40.0%; Score 54; DB 2; Length 137;
Best Local Similarity 46.4%; Pred. No. 1.6;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK-----STVSHPDLPREVRSIAKC 25

RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC005467; AAH05467.1; -.

```

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N1961 / SEROTYPE O1;
RA TISSUE=LYMPH, LYMPHOMA;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC002963; AAH02963.1; -
DR HSSP: P01825; 7FAB.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig-cl.
DR InterPro: IPR003600; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 5
DR SMART: SM00409; Ig; 2.
DR SMART: SM00407; Igcl; 4.
DR SMART: SM00406; IgV; 1.
DR SMART: SM00410; Ig_Like; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65274 MW; 2DFA8FB7E055851 CRC64;

Query Match 37.8%; Score 51; DB 4; Length 597;
Best Local Similarity 39.1%; Pred. No. 22;
Matches 9; Conservative 7; Mismatches 7; Indels 7; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRISIAK 24
|| : ||: ||| : ||:
Db 442 GERFTCTVTHDPLSPKQTISR 464

RESULT 13
Q96BB9 ID Q96BB9 PRELIMINARY; PRT; 597 AA.
AC Q96BB9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 65.0 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC015760; AAH15760.1; -
KW Hypothetical protein.
SQ SEQUENCE 597 AA; 65039 MW; 4FCA3AD8ECE263D9 CRC64;

Query Match 37.8%; Score 51; DB 4; Length 597;
Best Local Similarity 39.1%; Pred. No. 22;
Matches 9; Conservative 7; Mismatches 7; Indels 7; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRISIAK 24
|| : ||: ||| : ||:
Db 442 GERFTCTVTHDPLSPKQTISR 464

RESULT 14
Q9KPB0 ID Q9KPB0 PRELIMINARY; PRT; 597 AA.
AC Q9KPB0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GTP-BINDING PROTEIN LEPA.
GN VC2463.
OS Vibrio cholerae.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N1961 / SEROTYPE O1;
RA MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae.";
RL Nature 406:477-483(2000).
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL: AE004316; AAF95605.1; -
DR HSSP: P13551; 1DAR.
DR TIGR: VC2463; -
DR InterPro: IPR002106; AA_tRNA_ligase_II.
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR001806; Ras_trnsmng.
DR PRINTS: PR00315; ELONGATNFCT.
DR PRINTS: PR00449; RASTRNSPRMG.
DR PROSITE: PS00339; AA_tRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE: PS00301; EFACTOR_GTP; 1.
KW Complete proteome; GTP-binding.
SQ SEQUENCE 597 AA; 66029 MW; 754C1EE2E3DF0DDE CRC64;

Query Match 37.8%; Score 51; DB 16; Length 597;
Best Local Similarity 42.9%; Pred. No. 22;
Matches 12; Conservative 3; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHDPLPREVVRISIAK 25
||:| : | : | : | : | : |
Db 68 GQTYQLNFIDTPGHVDFAYEVSRLAAC 95

RESULT 15
Q96EY0 ID Q96EY0 PRELIMINARY; PRT; 613 AA.
AC Q96EY0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:20337).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC011857; AAH11857.1; -
SQ SEQUENCE 613 AA; 67273 MW; 31214203FB8421E7 CRC64;

Query Match 37.8%; Score 51; DB 4; Length 613;
Best Local Similarity 39.1%; Pred. No. 23;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRISIAK 24
|| : ||: ||| : ||:
Db 437 GERFTCTVTHDPLSPKQTISR 459

Search completed: August 26, 2002, 10:16:22
Job time: 713 sec
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:07:02 ; Search time 82.75 Seconds
(Without alignments)
33.557 Million cell updates/sec

Title: US-09-701-623C-84

Perfect score: 135

Sequence: 1 CGETKSTVSHPDLPREVRSIAKC 25

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_032802.*

1:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135	100.0	25	21	Optimised IgE-CH3
2	105	77.8	25	21	Optimised IgE-CH3
3	105	77.8	45	21	IgE immunogenic pe
4	105	77.8	45	21	IgE immunogenic pe
5	105	77.8	46	21	IgE immunogenic pe
6	105	77.8	57	21	IgE immunogenic pe
7	105	77.8	62	21	IgE immunogenic pe
8	105	77.8	63	21	IgE immunogenic pe
9	94	69.6	25	21	Optimised IgE-CH3
10	94	69.6	25	21	Modified human IgE
11	94	69.6	25	21	Peptide sequence o

12	94	69.6	42	21	AA1980014	IgE immunogenic pe
13	94	69.6	42	21	AA191215	Modified MVF Th ep
14	94	69.6	42	21	AA191216	Modified MVF Th ep
15	94	69.6	42	21	AA191217	Modified MVF Th ep
16	94	69.6	42	21	AA191218	Peptide sequence o
17	94	69.6	45	21	AA191219	IgE-CH3 domain ant
18	94	69.6	45	21	AA191220	Modified HBV surfa
19	94	69.6	45	21	AA191221	Peptide sequence o
20	94	69.6	46	21	AA191222	IgE immunogenic pe
21	94	69.6	46	21	AA191223	Modified MVF Th ep
22	94	69.6	46	21	AA191224	Modified MVF Th ep
23	94	69.6	46	21	AA191225	Peptide sequence o
24	94	69.6	56	21	AA191226	IgE immunogenic pe
25	94	69.6	59	21	AA191227	IgE immunogenic pe
26	94	69.6	60	21	AA191228	IgE immunogenic pe
27	94	69.6	60	21	AA191229	IgE immunogenic pe
28	94	69.6	63	21	AA191230	IgE-CH3 domain ant
29	94	69.6	63	21	AA191231	IgE immunogenic pe
30	94	69.6	63	21	AA191232	Inv epitope/modifi
31	94	69.6	63	21	AA191233	Peptide sequence o
32	89	65.9	25	21	AA191234	Optimised IgE-CH3
33	89	65.9	45	21	AA191235	IgE immunogenic pe
34	89	65.9	46	21	AA191236	IgE immunogenic pe
35	87	64.4	60	21	AA191237	IgE immunogenic pe
36	82	60.7	124	18	AA191238	Partial canine imm
37	82	60.7	312	21	AA191239	Dog immunoglobulin
38	82	60.7	341	21	AA191240	Immunogenic peptid
39	82	60.7	417	18	AA191241	Canine IgE heavy c
40	82	60.7	426	17	AA191242	Canine IgE. Canis
41	81	60.0	345	21	AA191243	Immunogenic peptid
42	74	54.8	25	21	AA191244	Optimised IgE-CH3
43	72	53.3	110	14	AA191245	Variant IgE - muta
44	71	52.6	106	20	AA191246	Human IgE Fc epsilon
45	71	52.6	110	14	AA191247	IgE Fc epsilon 3.

ALIGNMENTS

RESULT 1

AA1980077

ID AA1980077 standard; Peptide; 25 AA.

XX AC AA1980077;

XX DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
antibody; allergy; allergic disease; immunisation; anti-allergic;
anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Equus caballus.

XX OS Synthetic.

PN WO9967293-Al.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI: 2000-160578/14.

XX DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PT for immunization against allergy

The present invention describes immunoglobulin E (IgE)-CH3 domain antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic and anti-asthmatic properties. (I) induces polyclonal antibodies specific for a target effector site on the epsilon-heavy chain of IgE, and so preventing triggering and activation of mast cells and basophils and downregulation of IgE synthesis. Conjugates, or fusion peptides, containing (I) are used for active immunisation against IgE-mediated allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy dermatitis. Nucleic acids that encode these compounds are useful for recombinant production of corresponding peptides or in DNA vaccines. Conjugates of (I) that include a promiscuous T helper cell epitope (functional in genetically diverse subjects), in addition to a B cell target epitope, have increased immunogenicity and may include cyclic constraints (disulfide bridge) to stabilise conformational features and maximize cross-reactivity to the natural target. They induce safe (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino acid sequences used in the exemplification of the present invention.

1000

XX

DE IgE immunogenic peptide conjugate SEQ ID NO:91.
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
OS Unidentified.
XX
XX WO9967293-A1.
PN
XX 29-DEC-1999.
PD
XX
XX 21-JUN-1999; 99WO-US13959.
PF
XX
XX 20-JUN-1998; 98US-0100287.
PR
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
PA
XX Wang CY, Walfield AM;
PI
XX WPI; 2000-160578/14.
DR
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
PT
XX
XX Claim 14; Page 77; 155pp; English.
PS
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
XX Sequence 63 AA;
SQ
Query Match 77.8%; Score 105; DB 21; Length 63;
Best Local Similarity 72.0%; Pred. No. 3.4e-09;
Matches 18; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 39 cgetysrvthphlphkdivrsiack 63
RESULT 9
AAY79998
ID AAY79998 standard; Peptide; 25 AA.
XX
XX AAY79998;
AC
XX 15-MAY-2000 (first entry)
DT
XX Optimised IgE-CH3 domain antigen peptide for human IgE.
DE
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX

OS Homo sapiens.
OS Synthetic.
XX
XX WO9967293-A1.
PN
XX 29-DEC-1999.
PD
XX
XX 21-JUN-1999; 99WO-US13959.
PF
XX
XX 20-JUN-1998; 98US-0100287.
PR
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
PA
XX Wang CY, Walfield AM;
PI
XX WPI; 2000-160578/14.
DR
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
PT
XX
XX Claim 1; Page 21; 155pp; English.
PS
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
XX Sequence 25 AA;
SQ
Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 6.1e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 1 cgetysrvthphlpralmrsttkc 25
RESULT 10
AAY91212
ID AAY91212 standard; peptide; 25 AA.
XX
XX AAY91212;
AC
XX 22-MAY-2000 (first entry)
DT
XX Modified human IgE CH3 domain, SEQ ID NO:92.
DE
XX Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
XX Homo sapiens.
OS
OS Synthetic.
XX

PN	WO9966957-A2.	Query Match	69.6%;	Score 94;	DB 21;	Length 25;
XX	29-DEC-1999.	Best Local Similarity	64.0%;	Pred. No. 6.1e-08;		
XX		Matches 16;	Conservative	4;	Mismatches 5;	Indels 0;
XX	21-JUN-1999;					
XX	20-JUN-1998;					
XX	98US-0100412.					
XX	(UNBI-) UNITED BIOMEDICAL INC.					
XX	Wang CY;					
XX	WPI; 2000-160564/14.					
DR						
XX	New artificial T helper cell epitope and derived immunogens with target					
XX	antigenic site, for immunization against e.g. malaria, arteriosclerosis					
XX	or human immune deficiency virus -					
PS	Example 6; Page 40; 129pp; English.					
XX	The invention relates to novel promiscuous T helper cell epitopes (Th),					
CC	and immunogenic peptides comprising the Th epitopes of the invention					
CC	along with B cell epitopes. The Th epitopes and peptide immunogens					
CC	containing them, are used to induce a T helper cell response,					
CC	specifically against Plasmodium falciparum, cholesterol ester transport					
CC	protein (CETP) or HIV epitopes, but more generally against any pathogen,					
CC	immunoreactive self-antigen or tumour antigen. The Th epitopes and					
CC	peptide immunogens may be used for prevention and/or treatment of					
CC	infections (HIV, foot-and-mouth disease or malaria); for cancer					
CC	immunotherapy; for inhibition of the action of luteinising hormone					
CC	releasing hormone (LHRH) for contraception, treatment of hormone-					
CC	dependent cancer, prevention of boar taint in meat, and					
CC	immunocastration); for promoting the growth of animals; or for					
CC	treating allergies or arteriosclerosis. Incorporation of a promiscuous					
CC	Th (functional in genetically diverse subjects) into an immunogen					
CC	improves capacity to induce a strong T helper cell-mediated immune					
CC	response, resulting in production of antibodies against a target					
CC	antigen. Th can replace carrier proteins and pathogen-derived T helper					
CC	epitopes. Sequence AAY91121 represents a promiscuous T helper epitope					
CC	from the measles virus F (MVF) protein and sequences AAY91122-Y91142,					
CC	AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the					
CC	MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope					
CC	from hepatitis B virus (HBV) surface antigen, and sequences					
CC	AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.					
CC	AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides					
CC	comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197					
CC	is the LHRH target antigenic peptide used in these LHRH antigenic					
CC	peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic					
CC	peptides comprising somatostatin and a Th epitope. Somatostatin					
CC	immunogens may be used to promote growth in livestock. AAY91208 is a					
CC	human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th					
CC	epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV					
CC	infection of T cells. AAY90212 is a modified version of a human IgE					
CC	(immunoglobulin E) C _H 3 domain, and AAY90213-Y90219 are Th epitope/IgE C _H 3					
CC	antigenic peptides which may be used in the treatment of allergies.					
CC	AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)					
CC	VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th					
CC	epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target					
CC	antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th					
CC	epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent					
CC	CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a					
CC	CETP peptide and a Th epitope which may be used to prevent or treat					
CC	arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257					
CC	are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and					
CC	AAY91258-Y91273 are antigenic peptides, comprising MVH Th and HIV-1 B-cell					
CC	epitope which may be used as a component in an anti-HIV-1 vaccine.					
CC	AAY91198 and AAY91199 are respectively an immunostimulatory invasive					
CC	protein epitope from <i>Yersinia</i> species, and hinge spacer peptide, both of					
CC	which may optionally be used in the antigenic peptides of the					
XX	invention.					
XX	Sequence 25 AA;					
SQ						
PN	WO9966957-A2.	Query Match	69.6%;	Score 94;	DB 21;	Length 25;
XX	29-DEC-1999.	Best Local Similarity	64.0%;	Pred. No. 6.1e-08;		
XX		Matches 16;	Conservative	4;	Mismatches 5;	Indels 0;
XX	21-JUN-1999;					
XX	20-JUN-1998;					
XX	98US-0100412.					
XX	(UNBI-) UNITED BIOMEDICAL INC.					
XX	Wang CY;					
XX	WPI; 2000-160562/14.					
DR						
XX	New peptide immunogen containing luteinising hormone-releasing hormone					
XX	antigen site and helper T cell epitope, for e.g. contraception and					
XX	treatment of cancer -					
XX	Disclosure; Page 92; 102pp; English.					
XX	The specification describes peptide immunogens comprising a					
XX	synthetic helper T cell (Th) epitope and a target antigen, luteinising					
XX	hormone-releasing hormone (LHRH). The peptide immunogens cause					
XX	induction of a specific immune response to LHRH which is involved in					
XX	regulation of spermatogenesis, ovulation, oestrus, sexual development					
XX	and secretion of sex hormones. Provision of a promiscuous T helper					
XX	epitope (which is functional in genetically diverse subjects) provides					
XX	optimum immunogenicity to the B cell epitopes of the target antigen and					
XX	thus high antibody titres against the target antigen. The peptide					
XX	immunogens of the invention are used to vaccinate against mammalian LHRH,					
XX	for use as (reversible) contraceptive; control of hormone-dependent					
XX	tumours (cancer of prostate or breast, also endometriosis); to prevent					
XX	boar taint (and improve meat quality) and for immunocastration. The					
XX	present sequence appears in the specification.					
XX	Sequence 25 AA;					
SQ						
PN	WO9966957-A2.	Query Match	69.6%;	Score 94;	DB 21;	Length 25;
XX	29-DEC-1999.	Best Local Similarity	64.0%;	Pred. No. 6.1e-08;		
XX		Matches 16;	Conservative	4;	Mismatches 5;	Indels 0;
XX	21-JUN-1999;					
XX	20-JUN-1998;					
XX	98US-0100414.					
XX	(UNBI-) UNITED BIOMEDICAL INC.					
XX	Wang CY;					
XX	WPI; 2000-160562/14.					
DR						
XX	New peptide immunogen containing luteinising hormone-releasing hormone					
XX	antigen site and helper T cell epitope, for e.g. contraception and					
XX	treatment of cancer -					
XX	Disclosure; Page 92; 102pp; English.					
XX	The specification describes peptide immunogens comprising a					
XX	synthetic helper T cell (Th) epitope and a target antigen, luteinising					
XX	hormone-releasing hormone (LHRH). The peptide immunogens cause					
XX	induction of a specific immune response to LHRH which is involved in					
XX	regulation of spermatogenesis, ovulation, oestrus, sexual development					
XX	and secretion of sex hormones. Provision of a promiscuous T helper					
XX	epitope (which is functional in genetically diverse subjects) provides					
XX	optimum immunogenicity to the B cell epitopes of the target antigen and					
XX	thus high antibody titres against the target antigen. The peptide					
XX	immunogens of the invention are used to vaccinate against mammalian LHRH,					
XX	for use as (reversible) contraceptive; control of hormone-dependent					
XX	tumours (cancer of prostate or breast, also endometriosis); to prevent					
XX	boar taint (and improve meat quality) and for immunocastration. The					
XX	present sequence appears in the specification.					
XX	Sequence 25 AA;					
SQ						

```

RESULT 12
AAV80014
ID AAY80014 standard; Peptide; 42 AA.
XX AC AAY80014;
XX AC AAY80014;
DT 15-MAY-2000 (first entry)
XX DE IgE immunogenic peptide conjugate SEQ ID NO: 21.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Unidentified.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PT for immunization against allergy -
XX PS Claim 14; Page 76; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies
XX CC specific for a target effector site on the epsilon-heavy chain of IgE,
XX CC and so preventing triggering and activation of mast cells and basophils
XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe
XX CC (non-anaphylactogenic) antibodies. AAV79994 to AAY80084 represent amino
XX CC acid sequences used in the exemplification of the present invention.
XX SQ Sequence 42 AA;

Query Match 69.6%; Score 94; DB 21; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.2e-07;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETKSTVSHPDLPREVVRISARC 25
||||| 1:11111111111111111111
Db 18 cgetyqsrvtphlpralmrsttkc 42

RESULT 13
AAV91215
ID AAY91215 standard; peptide; 42 AA.
XX AC AAY91215;

```

```

XX 22-MAY-2000 (first entry)
XX Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:95.
XX Promiscuous T-cell epitope; measles virus F protein; MVP;
XX hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
XX somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
XX foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
XX Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
XX cholesterol ester transport protein; anti-arteriosclerotic.
XX Chimeric - Measles virus.
XX OS Chimeric - Homo sapiens.
XX PN WO9966957-A2.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13975.
XX PR 20-JUN-1998; 98US-0100412.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX DR WPI; 2000-160564/14.
XX PT New artificial T helper cell epitope and derived immunogens with target
XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX PT or human immune deficiency virus -
XX PS Example 6; Page 98-99; 129pp; English.
XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
XX CC and immunogenic peptides comprising the Th epitopes and peptide immunogens
XX CC along with B cell epitopes. The Th epitopes and peptide immunogens
XX CC containing them, are used to induce a T helper cell response,
XX CC specifically against Plasmodium falciparum, cholesterol ester transport
XX CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,
XX CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX CC peptide immunogens may be used for prevention and/or treatment of
XX CC infections (HIV, foot-and-mouth disease or malaria); for cancer
XX CC immunotherapy; for inhibition of the action of luteinising hormone
XX CC releasing hormone (LHRH) for contraception, treatment of hormone-
XX CC dependent cancer, prevention of boar taint in meat, and
XX CC immunocastration; for promoting the growth of animals; or for
XX CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX CC Th (functional in genetically diverse subjects) into an immunogen
XX CC improves capacity to induce a strong T helper cell-mediated immune
XX CC response, resulting in production of antibodies against a target
XX CC antigen. Th can replace carrier proteins and pathogen-derived T helper
XX CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
XX CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
XX CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
XX CC from hepatitis B virus (HBV) surface antigen, and sequences
XX CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
XX CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
XX CC is the LHRH target antigenic peptide used in these LHRH antigenic
XX CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
XX CC peptides comprising somatostatin and a Th epitope. Somatostatin
XX CC immunogens may be used to promote growth in livestock. AAY91208 is a
XX CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y91211 are MVF Th
XX CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
XX CC infection of T cells. AAY90212 is a modified version of a human IgE
XX CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
XX CC antigenic peptides which may be used in the treatment of allergies.
XX CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
XX CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th

```


KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX Chimeric - Measles virus.
OS Chimeric - Homo sapiens.
XX
XX WO9966957-A2.
PN
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX
PS Example 6; Page 99; 129pp; English.
XX
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CTP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone-
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CTP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CTP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasion
CC protein epitope from Yersinia species, and hinge spacer peptide, both of

CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
SQ Sequence 42 AA;

Query Match 69.6%; Score 94; DB 21; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.2e-07;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYSKSTVSHPDLPREVVRISIAKC 25
|||||:|:|:|:|:|:|:|:|:|:|
Db 18 cgetysgrvthphlpralmrsttkc 42

Search completed: August 26, 2002, 10:07:02
Job time: 228 sec

```

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No 6028987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

```

```

Query Match 69.6%; Score 94; DB 4; Length 25;
Best Local Similarity 64.0%; Pred. No. 8e-09;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 CGETKSTVSHDPDLPREVVRSTAKC 25
|||||:|:|:|:|:|:|:|
DB 1 CGETYSRVTHPHLPALMRSTTKC 25

```

```

RESULT 3
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows

```

```

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

```

```

Query Match 69.6%; Score 94; DB 3; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.5e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 CGETKSTVSHDPDLPREVVRSTAKC 25
|||||:|:|:|:|:|:|:|
DB 18 CGETYSRVTHPHLPALMRSTTKC 42

```

```

RESULT 4
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

```

```

Query Match 69.6%; Score 94; DB 3; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.5e-08;

```



```

:
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC Windows
: SOFTWARE: Word 97
:
:

```

```

1
2 APPLICATION NUMBER: US/09/303,323
3 FILING DATE: 30-APR-1999
4
5 CLASSIFICATION:
6
7 PRIOR APPLICATION DATA:
8 APPLICATION NUMBER: 09/100,414
9 FILING DATE: 20-JUNE-1998
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Maria H. Lin
12 REGISTRATION NUMBER: 29,323
13 REFERENCE/DOCKET NUMBER: 1151-4157
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: 212-758-4800
16 TELEFAX: 212-751-6849
17 INFORMATION FOR SEQ ID NO: 98:
18 SEQUENCE CHARACTERISTICS:
19 LENGTH: 42 amino acids
20 TYPE: amino acid
21 TOPOLOGY: linear
22 MOLECULE TYPE: peptide
23
24 US-09-303-323-98

```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match          69.6%      Score 94;   DB 4;   Length 42;
Best Local Similarity 64.0%      Pred. No. 1.5e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

ADDRESS: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999

```

/ OPERATING SYSTEM: PC Windows
/ SOFTWARE: Word 97
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/303,323
/ FILING DATE: 30-APR-1999
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/100,414
/ FILING DATE: 20-JUNE-1998
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Maria H. Lin
/ REGISTRATION NUMBER: 29,323
/ REFERENCE/DOCKET NUMBER: 1151-4157
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-758-4800
/ TELEFAX: 212-751-6849
/ INFORMATION FOR SEO ID NO: 99

```

RESULT • 9
US-09-100-414B-101
: Sequence 101, Application US/09100414B

```

RESULT 10
US-09-303-323-101
; Sequence 101, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

```
,  
STATE: NI  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible
```

OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-96

Query Match 69.6%; Score 94; DB 4; Length 46;
Best Local Similarity 64.0%; Pred. No. 1.7e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 22 CGETYOSRVTHPLPRALMRSTTKC 46

RESULT 14

US-09-303-323-97
Sequence 97, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-97

Query Match 69.6%; Score 94; DB 4; Length 46;
Best Local Similarity 64.0%; Pred. No. 1.7e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 22 CGETYOSRVTHPLPRALMRSTTKC 46

RESULT 15

US-09-100-414B-102
Sequence 102, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-102

Query Match 69.6%; Score 94; DB 3; Length 63;
Best Local Similarity 64.0%; Pred. No. 2.5e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 39 CGETYOSRVTHPLPRALMRSTTKC 63

Search completed: August 26, 2002, 10:05:19
Job time: 125 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:17:02 ; Search time 24.08 Seconds
(without alignments)
40.199 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPFPKPIVRITKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	122	84.1	429	1 EPC_RAT	P01855 rattus norv
2	98	67.6	421	1 EPC_MOUSE	P06336 mus musculus
3	80	55.2	428	1 EPC_HUMAN	P01854 homo sapien
4	51	35.2	264	1 HB2D_RAT	P18211 rattus norv
5	49	33.8	345	1 HALF_CHICK	P15979 gallus gall
6	48.5	33.4	704	1 KPCI_CAEEL	P34722 caenorhabdi
7	48	33.1	324	1 GCI_MOUSE	P01868 mus musculus
8	48	33.1	393	1 GCI_MOUSE	P01869 mus musculus
9	47	32.4	322	1 GCA_RAT	P20760 rattus norv
10	46.5	32.1	345	1 DER2_HUMAN	O9h3k2 homo sapien
11	46	31.7	116	1 SOR_METJA	Q58151 methanococc
12	46	31.7	277	1 VNST_CVEHQ	P18517 bovine coro
13	46	31.7	336	1 GCB_MOUSE	P01866 mus musculus
14	46	31.7	370	1 DFCUP_SCHPO	O9usj5 schizosacch
15	46	31.7	373	1 MTF2_MOUSE	O02395 mus musculus
16	46	31.7	405	1 GCB_MOUSE	P01867 mus musculus
17	46	31.7	516	1 MEFA_XENLA	Q03414 xenopus lae
18	46	31.7	593	1 MTF2_HUMAN	O9y483 homo sapien
19	45	31.0	116	1 YK1A_YEAST	P36078 saccharomyc
20	45	31.0	330	1 GCAA_MOUSE	P01863 mus musculus
21	45	31.0	399	1 GCM_MOUSE	P01865 mus musculus
22	44.5	30.7	387	1 GCP_RICPR	O9zeas8 rickettsia
23	44	30.3	104	1 Y4EB_RHLSN	P55425 rhizobium s
24	44	30.3	263	1 HB2B_RAT	P29826 rattus norv
25	44	30.3	275	1 Y760_METJA	O58170 methanococc
26	44	30.3	290	1 GC3_HUMAN	P01860 homo sapien
27	44	30.3	326	1 GC2_HUMAN	P01859 homo sapien
28	44	30.3	329	1 GC3_MOUSE	P22436 mus musculus
29	44	30.3	330	1 GCI_HUMAN	P01857 homo sapien
30	44	30.3	371	1 HA12_RAT	P16391 rattus norv
31	44	30.3	371	1 RIBD_BACAM	P70814 b riboflavi
32	44	30.3	384	1 GBA2_PEA	O04279 pisum sativ
33	44	30.3	398	1 GC3_MOUSE	P03987 mus musculus

34 44 30.3 446 1 MUC_CHICK P01875 gallus gall
35 44 30.3 1106 1 KPCI_LYTP1 Q25378 lytechinus
36 44 30.3 581 1 FZD8_XENLA P53585 caenorhabdi
37 43.5 30.0 581 1 OC90_CAVPO O33274 xenopus lae
38 43 29.7 51 1 RS20_HUMAN P18169 cavia porce
39 43 29.7 119 1 RS20_HUMAN P23403 xenopus lae
40 43 29.7 119 1 RS20_XENLA P34075 homo sapien
41 43 29.7 120 1 RS20_DROME P25828 drosophila
42 43 29.7 252 1 HB2F_MOUSE P06346 mus musculus
43 43 29.7 261 1 HB24_HUMAN P01920 homo sapien
44 43 29.7 263 1 HB2K_MOUSE P06343 mus musculus
45 43 29.7 263 1 HB2S_MOUSE P06345 mus musculus

ALIGNMENTS

RESULT 1

ID EPC_RAT STANDARD; PRT; 429 AA.
AC P01855;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC STRAIN=LOU/C/WSL;
RX MEDLINE=83084537; PubMed=6292865;
RA Hellman L., Pettersson U., Engstrom A., Karlsson T., Bennich H.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RT "Structure and evolution of the heavy chain from rat immunoglobulin
RT E.";
RL Nucleic Acids Res. 10:6041-6049(1982).
RN [2]
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6820340;
RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RT construction, identification, and DNA sequence.";
RL DNA 1:335-343(1982).
RN [3]
RP SEQUENCE OF 205-306 FROM N.A.
RX MEDLINE=82174576; PubMed=6803238;
RA Hellman L., Pettersson U., Bennich H.;
RT "Characterization and molecular cloning of the mRNA for the heavy
RT (epsilon) chain of rat immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1284-1268(1982).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00744; AAA41379.1; ALT_INIT.
DR PIR: A02143; EHRT.
DR HSPSP; P01854; LIGE.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00410; IG-like; 3.
DR SMART; SM00407; IG1; 1.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_REF 1 168 R -> N (IN REF. 2).
FT CONFLICT 168 168

FT CONFLICT 308 308 P -> L (IN REF. 2).
SQ SEQUENCE 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;

Query Match 84.1%; Score 122; DB 1; Length 429;
Best Local Similarity 95.7%; Pred. No. 1.6e-10;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYSRVDPHPKPIVRSITK 24

DB 284 GEGYQCRVDHPKPIVRSITK 306

RESULT 2

EPC_MOUSE STANDARD; PRT; 421 AA.

AC P06336; P01856;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Ig epsilon chain C region.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=84236092; PubMed=6329728;

RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;

RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:

comparison with the human epsilon gene sequence.";

RL EMBO J. 1:1117-1123(1982).

RN [2]

RP REVISIONS.

RA Honjo T.;

RL Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 34-421 FROM N.A.

RX MEDLINE=8311774; PubMed=6818553;

RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;

RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon

chain cDNA.";

RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X01857; CAA25977.1; -

DR EMBL; X01857; CAA25978.1; -

DR PIR; A02145; EHMS.

DR PIR; A02144; EHMS.

DR HSSP; P01854; 1IGE.

DR InterPro; IPR003006; Ig_MHC.

DR InterPro; IPR003597; Ig_C1.

DR InterPro; IPR003600; Ig_Like.

DR Pfam; PF00047; Ig_4.

DR SMART; SM00410; Ig_Like; 2.

DR SMART; SM00407; IGc1; 2.

DR PROSITE; PS00290; IG_MHC; 3.

KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON_TER 1 1

FT DOMAIN 1 90 CH1.

FT DOMAIN 91 197 CH2.

FT DOMAIN 198 304 CH3.

FT DOMAIN 305 421 CH4.

FT DISULFID 23 75 BY SIMILARITY.

FT DISULFID 121 180 BY SIMILARITY.

FT DISULFID 226 285 BY SIMILARITY.

FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909EIF30A06B47 CRC64;

Query Match 67.6%; Score 98; DB 1; Length 421;

Best Local Similarity 82.6%; Pred. No. 5.2e-07;

Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GEGYSRVDPHPKPIVRSITK 24

DB 280 GGYQCIVDPDPKPIVRSITK 302

RESULT 3

EPC_HUMAN

ID EPC_HUMAN

AC P01854;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ig epsilon chain C region.

GN IGHE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83168897; PubMed=6300763;

RA Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,

RA Kikuchi M., Sugino Y., Nishida Y., Honjo T.;

RT "Molecular cloning and nucleotide sequencing of human immunoglobulin

epsilon chain cDNA.";

RL Nucleic Acids Res. 11:719-726(1983).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=83001945; PubMed=6288268;

RA Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;

RT "Duplication and deletion in the human immunoglobulin epsilon genes.";

RL Cell 29:691-699(1982).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=84236029; PubMed=6234164;

RA Flanagan J.G., Rabbitts T.H.;

RT "The sequence of a human immunoglobulin epsilon heavy chain constant

region gene, and evidence for three non-allelic genes.";

RL EMBO J. 1:655-660(1982).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=84207910; PubMed=6327276;

RA Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;

RT "Long terminal repeat-like elements flank a human immunoglobulin

epsilon pseudogene that lacks introns.";

RL EMBO J. 1:1539-1544(1982).

RN [5]

RP PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).

RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;

RL (In) Bach M.K. (eds.);

RT Immediate hypersensitivity: modern concepts and developments, pp.1-36,

RL Marcel Dekker, New York (1978).

RN [6]

RP SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.

RX MEDLINE=83065234; PubMed=6815656;

RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,


```

RA Beil L.O., Gould H.J.;
RT "Cloning and sequence determination of the gene for the human
RL immunoglobulin epsilon chain expressed in a myeloma cell line.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
[7]
RA 3D-STRUCTURE MODELING.
RP Padlan E.A., Davies D.R.;
RT "A model of the F-C $\gamma$  of immunoglobulin E.";
RL Submitted (JUL-1993) to the PDB data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L00022; AAB59424.1; ALT_INIT.
DR PIR; A02142; EHHU.
DR PIR; A22771; A22771.
DR PIR; A23195; A23195.
DR PDB; 1IGE; 15-JUL-92.
DR MIM; 147180; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00407; IGC1; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure.
FT NON_TER 1 1 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 14 14
FT DISULFID 15 105
FT DISULFID 29 85
FT DISULFID 121 121 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 135 193
FT DISULFID 209 209 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 239 299
FT DISULFID 345 405
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .).
FT VARIANT 359 359 W -> L (IN REF. 2, POSSIBLY DUE TO
FT POLYMORPHISM).
FT SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;
/FTID=VAR_003885.

Query Match 55.2%; Score 80; DB 1; Length 428;
Best Local Similarity 60.9%; Pred. No. 0.00023;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
||| :||| :|||
Db 294 GETYQCRVTHPHLPALMRSTTK 316

RESULT 4
HB2D_RAT
ID HB2D_RAT STANDARD; PRT; 264 AA.
AC P18211;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE RT1 class II histocompatibility antigen, D-1 beta chain precursor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=LEWIS FAMILIARIS; TISSUE=Bone marrow;
RX MEDLINE=90356406; PubMed=2388838;
RA Syha-Jedelhauser J., Reske K.;
RT "Sequence of rat cDNA clone pLR beta 112 coding for the RT1.D beta I
RT chain.";
RL Nucleic Acids Res. 18:4598-4598(1990).
CC -|- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53054; CAA37221.1; -.
DR PIR; S10989; S10989.
DR HSSP; P13760; ZSEB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR PRODOM; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC II; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 264 RT1 CLASS II HISTOCOMPATIBILITY ANTIGEN,
FT D-1 BETA CHAIN.
FT DOMAIN 27 120 EXTRACELLULAR BETA-1.
FT DOMAIN 121 215 EXTRACELLULAR BETA-2.
FT DOMAIN 216 226 CONNECTING PEPTIDE.
FT TRANSMEM 227 248
FT DOMAIN 249 264 CYTOPLASMIC TAIL.
FT DISULFID 42 106 BY SIMILARITY.
FT DISULFID 144 200 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 264 AA; 29876 MW; 158F357355177DA1 CRC64;

Query Match 35.2%; Score 51; DB 1; Length 264;
Best Local Similarity 47.1%; Pred. No. 2.4;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
||| :||| :|||
Db 195 GEVYTCQVEHPSLPSPV 211

RESULT 5
HAIF_CHICK
ID HAIF_CHICK STANDARD; PRT; 345 AA.
AC P15979;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Class I histocompatibility antigen, F10 alpha chain precursor (B-F
DE histocompatibility F10 antigen) (B-F-beta-IV) (B12).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=B12; TISSUE=Liver;
RX MEDLINE=89030642; PubMed=3141149;
RX Guillemot F., Billault A., Pourquie O., Behar G., Chausse A.M.,

```

```

RA Zoorob R., Kreibich G., Auffray C.;
RT "A molecular map of the chicken major histocompatibility complex: the
RT class II beta genes are closely linked to the class I genes and the
RT nucleolar organizer.";
RL EMBO J. 7:2775-2785(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90316612; PubMed=2370087;
RA Kroemer G., Zoorob R., Auffray C.;
RT "Structure and expression of a chicken MHC class I gene.";
RL Immunogenetics 31:405-409(1990).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X12780; CAA31272.1; -
CC EMBL; M31012; AAA48947.1; -
CC PIR; S01172; HLCB84.
CC PIR; A45846; A45846.
CC HSSP; P03989; LHSA.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_cl.
CC InterPro; IPR001039; MHC_I.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00129; MHC_I; 1.
CC ProDom; PD000050; MHC_I; 1.
CC SMART; SM00407; Igc1; 1.
CC PROSITE; PS00290; Ig_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 345
FT ALPHA CHAIN
FT CLASS I HISTOCOMPATIBILITY ANTIGEN, F10
FT ALPHA CHAIN
FT DOMAIN 23 110
FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 111 201
FT EXTRACELLULAR ALPHA-2.
FT DOMAIN 202 292
FT EXTRACELLULAR ALPHA-3.
FT DOMAIN 293 301
FT CONNECTING PEPTIDE.
FT TRANSMEM 302 324
FT DOMAIN 325 345
FT CYTOPLASMIC.
FT DISULFID 121 183
FT BY SIMILARITY.
FT DISULFID 221 277
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 59 59
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107
FT CONFLICT 339 339
FT P -> PDREGGSSSST (IN REF. 2).
SQ SEQUENCE 345 AA; 38246 MW; 430DCCF8091B69A4 CRC64;

Query Match 33.8%; Score 49; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 6.2;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17
Db 272 GDKYQCRVHASLPQP 287

RESULT 6
KPC1_CAEEL
ID KPC1_CAEEL STANDARD; PRT; 704 AA.
AC F34722;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein kinase C-like 1 (EC 2.7.1.-) (PKC).
GN TPA-1.-

```

```

OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX MEDLINE=95387388; PubMed=7658466;
RA Sano T., Tabuse Y., Nishiwaki K., Miwa J.;
RT "The tpa-1 gene of Caenorhabditis elegans encodes two proteins
RT similar to Ca(2+)-independent protein kinase Cs: evidence by complete
RT genomic and complementary DNA sequences of the tpa-1 gene.";
RL J. Mol. Biol. 251:477-485(1995).
RN [2]
RP SEQUENCE OF 148-704 FROM N.A.
RX MEDLINE=89186920; PubMed=2538925;
RA Tabuse Y., Nishiwaki K., Miwa J.;
RT "Mutations in a protein kinase C homolog confer phorbol ester
RT resistance on Caenorhabditis elegans.";
RL Science 243:1713-1716(1989).
RN [3]
RP REVISIONS TO REF.2.
RA Miwa J.;
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
CC THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: TPA-1A (SHOWN HERE) AND TPA-1B;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D49525; BAA08470.1; -
CC EMBL; D49525; BAA08471.1; -
CC EMBL; D14815; BAA03556.1; -
CC HSSP; P28867; IPTO.
CC InterPro; IPR002219; DAG_PE-bind.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR000961; Pkinase_C.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00130; DAG_PE-bind; 2.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00433; pkinase_C; 1.
CC PRINTS; PR00008; DAGPEDOMAIN.
CC SMART; SM00109; CI; 2.
CC SMART; SM00133; S_TK_X; 1.
CC SMART; SM00220; S_TKC; 1.
CC PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
CC PROSITE; PS50081; DAG_PE_BIND_DOM_2; 2.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;
KW Phorbol-ester binding; Repeat; Alternative splicing; Phosphorylation.
FT DOMAIN 166 215
FT PHORBOL-ESTER AND DAG BINDING 2.
FT DOMAIN 238 287
FT PHORBOL-ESTER AND DAG BINDING 2.
FT DOMAIN 375 634
FT PROTEIN KINASE.
FT NP_BIND 381 389
FT ATP (BY SIMILARITY).
FT BINDING 404 404
FT ATP (BY SIMILARITY).
FT ACT_SITE 499 499
FT BY SIMILARITY.
FT MOD_RES 89 89
FT PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT MOD_RES 139 139
FT PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT MOD_RES 324 324
FT PHOSPHORYLATION (AUTO-) (POTENTIAL).

```

```

FT VARSPLIC 1 137 MISSING (IN ISOFORM TPA-1B).
SQ SEQUENCE 704 AA; 80298 MW; C44F2E25F58057E3 CRC64;

Query Match 33.4%; Score 48.5; DB 1; Length 704;
Best Local Similarity 37.9%; Pred. No. 15;
Matches 11; Conservative 2; Mismatches 11; Indels 5; Gaps 1;

QY 2 GEGYSRVD-----HPHPKPPIVRSITK 25
||| | 1:1111: ||
Db 577 GEGELEDSDILNERPYFPKTSKEAAK 605

RESULT 7
GCLM_MOUSE STANDARD: PRT; 324 AA.
ID GCLM_MOUSE
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE=80202559; PubMed=6769752;
RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
cloned in a bacterial plasmid.";
RL Gene 9:87-97(1980).
RN [3]
RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=80012837; PubMed=113776;
RA Rogers J., Clarke P., Salsar W.;
RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
heavy chain.";
RL Nucleic Acids Res. 6:3305-3321(1979).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE=78242288; PubMed=98524;
RA Adeugbo K.;
RT "Evolution of immunoglobulin subclasses. Primary structure of a
murine myeloma gammal chain.";
RL J. Biol. Chem. 253:6068-6075(1978).
RN [5]
RP DISULFIDE BONDS (MOPC 21).
RX MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulphide bridges of a mouse immunoglobulin G1 proteinin.";
RL Biochem. J. 126:837-850(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -

```

```

DR EMBL; V00793; CAA24175.1; -
DR EMBL; V00795; CAA24176.1; -
DR PIR; A02159; GIMS.
DR HSSP; P01842; 7FAB.
DR GlycoSuiteDB; P01868; -
DR MGD; MGI:96446; Igh-4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
FT NON_TER 1
FT DOMAIN 1 97
FT HINGE.
FT DOMAIN 98 110
FT CH2.
FT DOMAIN 111 217
FT CH3.
FT DISULFID 27 82
FT DISULFID 102 102
FT DISULFID 104 104
FT DISULFID 107 107
FT DISULFID 109 109
FT DISULFID 138 198
FT CARBOHYD 174 174
FT N-LINKED (GLCNAC. . .).
FT /FTIG-CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324
FT CONFLICT 276 276
FT CONFLICT 278 278
FT SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;

Query Match 33.1%; Score 48; DB 1; Length 324;
Best Local Similarity 39.1%; Pred. No. 8.1;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYSRVDHPHPKPPIVRSITK 24
|: :|: ||| :||:|
Db 193 GKEFKCRVNSAAPPAPIETISK 215

RESULT 8
GCLM_MOUSE STANDARD: PRT; 393 AA.
ID GCLM_MOUSE
AC P01869;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80045036; PubMed=115593;
RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.;
RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
gamma 1 chain gene.";
RL Cell 18:559-568(1979).
RN [2]
RP SEQUENCE OF 323-393 FROM N.A.
RX MEDLINE=82197626; PubMed=6804950;
RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
RT "mRNA for surface immunoglobulin gamma chains encodes a highly
conserved transmembrane sequence and a 28-residue intracellular
domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
RN [3]
RP SEQUENCE OF 323-366 FROM N.A.
RX MEDLINE=82115295; PubMed=6799207;

```



```

RN [3]*
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mao Y.M., Xie Y., Mu Z.M., Li Y., Huang Y.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansgore W., Boecher M., Bloecher H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mexes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RT Genome Res. 11:422-435(2001).
RL
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE BIL FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009685; BAA33049.1; -
DR EMBL; AF131820; AAD20052.1; -
DR EMBL; AF060923; AAG43135.1; -
DR EMBL; AL136713; CAB66648.1; -
KW Transmembrane.
FT TRANSMEM 83 103 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 160 180 POTENTIAL.
FT TRANSMEM 191 211 POTENTIAL.
FT TRANSMEM 214 234 POTENTIAL.
FT TRANSMEM 245 265 POTENTIAL.
FT TRANSMEM 272 292 POTENTIAL.
FT CONFLICT 4 4 A -> V (IN REF. 4).
FT CONFLICT 64 64 E -> G (IN REF. 3).
FT CONFLICT 74 74 I -> M (IN REF. 4).
FT CONFLICT 118 118 Q -> R (IN REF. 4).
SQ SEQUENCE 345 AA; 37205 MW; 808FAED86A9CD98E CRC64;

Query Match 32.1%; Score 46.5; DB 1; Length 345;
Best Local Similarity 51.9%; Pred. No. 14;
Matches 14; Conservative 2; Mismatches 8; Indels 3; Gaps 2;

QY 1 CEGYQSRVDHPHPK--PIVR-SITK 24
| | | | | | | | | |
DB 8 CLRTLPSRVFHPFAFTKASPVVKNSTK 34

RESULT 11
SOR_METJA
ID SOR_METJA STANDARD; PRT; 116 AA.
AC Q58151;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative superoxide reductase (EC 1.18.96.1) (SOR).
GN MJ0741.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8668087;

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RT Science 273:1058-1073(1996).
RL
CC -!- FUNCTION: USES ELECTRONS FROM REDUCED NADP, BY WAY OF RUBREDOXIN
CC AND AN OXIDOREDUCTASE, TO CATALYZE THE REDUCTION OF SUPEROXIDE TO
CC HYDROGEN PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Reduced rubredoxin + superoxide + 2 H(+) =
CC H(2)O(2) + rubredoxin.
CC -!- COFACTOR: IRON (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE DESULFOFERRODOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67520; AAB98735.1; -
DR HSSP; P22076; 1DFX.
DR TIGR; MJ0741; -
DR InterPro: IPR002742; Desulfoferrodox.
DR Pfam: PF01880; Desulfoferrodox; 1.
DR ProDom: PD006618; Desulfoferrodox; 1.
KW Hypothetical protein; Oxidoreductase; Electron transport; Iron;
KW Complete proteome.
FT METAL 20 20 IRON (BY SIMILARITY).
FT METAL 46 46 IRON (BY SIMILARITY).
FT METAL 52 52 IRON (BY SIMILARITY).
FT METAL 101 101 IRON (BY SIMILARITY).
FT METAL 104 104 IRON (BY SIMILARITY).
SQ SEQUENCE 116 AA; 13950 MW; 4CF2C76237DE0673 CRC64;

Query Match 31.7%; Score 46; DB 1; Length 116;
Best Local Similarity 43.5%; Pred. No. 5.5;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
| | | | | | | | | |
DB 62 GDLYLARVDFTQFMKPEVKLMVK 84

RESULT 12
VNST_CVBQ
ID VNST_CVBQ STANDARD; PRT; 277 AA.
AC P18517;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 32 kDa nonstructural protein.
DE 32 kDa nonstructural protein.
OS Bovine coronavirus (strain Quebec).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=111133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345182; PubMed=2762160;
RA Cox G.J., Parker M.D., Babiuk L.A.;
RT "The sequence of cDNA of bovine coronavirus 32K nonstructural gene.";
RL Nucleic Acids Res. 17:5847-5847(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X15445; CAA33485.1; .

DR PIR; A34039; MNH32.

KW Nonstructural protein.

SQ SEQUENCE 277 AA; 31896 MW; D2EEC7BAFB86EA11 CRC64;

Query Match 31.7%; Score 46; DB 1; Length 277;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPH 13
||||:|:|:|

DB 78 GEGFQIKFDNPH 89

RESULT 13

GCB_MOUSE STANDARD; PRT; 336 AA.

ID GCB_MOUSE P01866;

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ig gamma-2b chain C region.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A. (A ALLELE).

RX MEDLINE=80120716; PubMed=6766534;

RA Yamawaki-Kataoka Y., Kataoka T., Takahashi N., Obata M., Honjo T.;

RT "Complete nucleotide sequence of immunoglobulin gamma2b chain gene

cloned from newborn mouse DNA.";

RL Nature 283:786-789(1980).

RN [2]

RP SEQUENCE FROM N.A. (MPC 11).

RX MEDLINE=80081501; PubMed=117548;

RA Tucker P.W., Marcu K.B., Slightom J.L., Blattner F.R.;

RT "Structure of the constant and 3' untranslated regions of the murine

gamma 2b heavy chain messenger RNA.";

RL Science 206:1299-1303(1979).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=80081502; PubMed=117549;

RA Tucker P.W., Marcu K.B., Newell N., Richards J., Blattner F.R.;

RT "Sequence of the cloned gene for the constant region of murine gamma

2b immunoglobulin heavy chain.";

RL Science 206:1303-1306(1979).

RN [4]

RP SEQUENCE FROM N.A. (B ALLELE).

RX MEDLINE=82173203; PubMed=6803173;

RA Ollo R., Rougeon F.;

RT "Mouse immunoglobulin allotypes: post-duplication divergence of gamma

2a and gamma 2b chain genes.";

RL Nature 296:761-763(1982).

RN [5]

RP CARBOHYDRATE-LINKAGE SITE THR-105.

RX MEDLINE=94216359; PubMed=7512967;

RA Kim H., Yamaguchi Y., Masuda K., Matsunaga C., Yamamoto K.,

RA Irimura T., Takahashi N., Kato K., Arata Y.;

RT "O-glycosylation in hinge region of mouse immunoglobulin G2b.";

RL J. Biol. Chem. 269:12345-12350(1994).

CC 1- PTM: O-LINKED GLYCAN CONSISTS OF GAL-GALNAc DISACCHARIDE WHICH IS

MODIFIED WITH 2 SIALIC ACID RESIDUES.

CC 4- MISCELLANEOUS: DISULFIDE BONDS BY HOMOLOGY WITH OTHER IG GAMMA

CHAINS.

CC 1- MISCELLANEOUS: THE A ALLELE SEQUENCE IS SHOWN.

PIR; A02157; G2MS11.
DR HSSP; P01857; 1FC1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00410; Ig_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON_TER 1

FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).

FT DISULFID 27 82

FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).

FT DISULFID 150 210

FT DISULFID 256 314

FT CARBOHYD 105 105

FT MOD_RES 336 336

FT VARIANT 163 163

FT VARIANT 194 194

FT VARIANT 300 300

FT VARIANT 301 301

FT CONFLICT 25 25

FT CONFLICT 36 36

FT CONFLICT 239 239

SQ SEQUENCE 336 AA; 36658 MW; 7D879662607C356E CRC64;

Query Match 31.7%; Score 46; DB 1; Length 336;

Best Local Similarity 34.8%; Pred. No. 17;

Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24

DB 205 GKEFKCKVNKDLSPSIERTISK 227

RESULT 14

DCUP_SCHPO

ID DCUP_SCHPO STANDARD; PRT; 370 AA.

AC Q90SJ5;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).

GN SPCC4B3.05C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA Seeger K., Harris D., McDougall R.C., Rajandream M.A., Barrell B.G.;

RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.

CC 1- CATALYTIC ACTIVITY: Uroporphyrinogen-III - coproporphyrinogen + 4

CO(2).

CC 1- PATHWAY: PORPHYRIN AND HEME BIOSYNTHESIS.

CC 1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC 1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

DR EMBL; AL132870; CAB60679.1; -
DR HSP; P06132; IURO.
DR InterPro; IPR000257; Uroporphyrnngn_decarbxylys.
DR Pfam; PF01208; URO-D; 1.
DR PROSITE; PS00906; UROD_1; FALSE_NEG.
DR PROSITE; PS00907; UROD_2; 1.
KW Lyase; Decarboxylase; Porphyrin biosynthesis; Heme biosynthesis.
SQ SEQUENCE 370 AA; 41846 MW; 110DFC463733FD89 CRC64;

Query Match 31.7%; Score 46; DB 1; Length 370;
Best Local Similarity 87.5%; Pred. No. 18;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHFKPKIV 19
IIIIIIII
DB 110 PHFKPKLV 117

Search completed: August 26, 2002, 10:17:03
Job time: 699 sec

RESULT 15
MTF2_MOUSE
ID MTF2_MOUSE STANDARD; PRT; 373 AA.
AC Q02395;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metal-response element-binding transcription factor 2 (Zinc-regulated
factor 1) (Zirf1) (Metal-response element DNA-binding protein M96)
DE (Fragment).
DE MTF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoma;
RX MEDLINE=95290093; PubMed=7772254;
RA Inouye C., Remondelli P., Karin M., Elledge S.;
RT "Isolation of a cDNA encoding a metal response element binding protein
using a novel expression cloning procedure: the one hybrid system.";
RL DNA Cell Biol. 13:731-742(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=97234834; PubMed=9173905;
RA Remondelli P., Leone A.;
RT "Interactions of the zinc-regulated factor (Zirf1) with the mouse
metallothionein Ia promoter.";
RL Biochem. J. 323:79-85(1997).
CC -!- FUNCTION: BINDS TO THE METAL-REGULATING-ELEMENT (MRE) OF
METALLOTHIONEIN IA GENE PROMOTER. BINDING IS ZINC-DEPENDENT.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S78454; AAC34714.1; -
DR MGD; MGI:105050; Mtf2.
DR InterPro; IPR001965; PHD.
DR Pfam; PF00628; PHD; 2.
DR SMART; SM00249; PHD; 2.
KW DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KW Metal-binding.
FT NON_TER 1 1
FT ZN_FING 2 55 PHD-TYPE 1.
FT ZN_FING 101 153 PHD-TYPE 2.
FT DOMAIN 235 238 POLY-LYS.

SQ SEQUENCE 373 AA; 42003 MW; 50D2ECD13A5DE817 CRC64;

Query Match 31.7%; Score 46; DB 1; Length 373;
Best Local Similarity 33.3%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRISITK 24
II:III
DB 23 CGQGYHQLCHTTHIDSSVIDSDEK 46

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:16:18 ; Search time 66.82 Seconds
(without alignments)
64.724 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPKPIVRISITKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertibrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Result No.	Score	Query Match	Length	DB ID	Description
1	62	42.8	2421	5 Q9VM88	Q9vm88 drosophila
2	54	37.2	317	16 Q98DQ7	Q98dq7 rhizobium 1
3	54	37.2	648	5 Q9NKD7	Q9nkd7 drosophila
4	54	37.2	701	5 Q9VJU4	Q9vju4 drosophila
5	52	35.9	217	7 Q30827	Q30827 ovis aries
6	51	35.2	245	7 Q31270	Q31270 rattus norv
7	51	35.2	261	7 Q9TQA5	Q9tqa5 rattus norv
8	51	35.2	264	7 Q9TQA5	Q9tqa5 rattus norv
9	50	34.5	140	12 Q99A59	Q99a59 bovine vira
10	50	34.5	142	7 Q95536	Q95536 pongo pygma
11	50	34.5	314	11 Q9ESH2	Q9esh2 rattus norv
12	50	34.5	337	6 Q95M34	Q95m34 equus caball
13	50	34.5	357	10 Q947H5	Q947h5 nicotiana t
14	50	34.5	357	10 Q947H4	Q947h4 nicotiana t
15	50	34.5	539	5 Q965D2	Q965d2 trypanosoma
16	49	33.8	125	11 Q9CZ59	Q9cz59 mus musculus

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

RESULT	ID	Q9VM88	PRELIMINARY;	PRT;	2421 AA.
AC	Q9VM88;				
DT	01-MAY-2000	(TREMBLrel. 13, Created)			
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)			
DE	CG11321	PROTEIN.			
GN	CG11321				
OS	Drosophila melanogaster	(Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephyrdoidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	STRAIN=BERKELEY;				
RX	MEDLINE=20196006;	PubMed=10731132;			
RA	Adams M.D., Celiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt J.H., Nelson C.R., Miklos G.L.G.,				
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				

ALIGNMENTS

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003615; AAF52435.1; -;
 DR FlyBase; FBgn0031857; CG11321.
 DR InterPro; IPR002106; AA_TRNA_ligase_II.
 DR InterPro; IPR001876; Znf-RanBP.
 DR SMART; SM00184; Znf_ring.
 DR SMART; SM00547; Znf_RBZ; 1.
 DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
 SQ SEQUENCE 2421 AA; 269655 MW; 6F15EFD9128F13E0 CRC64;

Query Match 42.8%; Score 62; DB 5; Length 2421;
 Best Local Similarity 52.9%; Pred. No. 0.73; Mismatches 3; Indels 5; Gaps 0;
 Matches 9; Conservative 3;

QY 2 GEGYQSRVDHPHPKPI 18
 DB 211 GPGYMSQQHHPYPPV 227

RESULT 2

Q98DQ7 ID Q98DQ7 PRELIMINARY; PRT; 317 AA.
 AC Q98DQ7;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE MLR4594 PROTEIN.
 GN MLR4594.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003004; BAB51213.1; -;
 KW Complete proteome.
 SQ SEQUENCE 317 AA; 33121 MW; F7DE20BB4F661F32 CRC64;

Query Match 37.2%; Score 54; DB 16; Length 317;
 Best Local Similarity 38.5%; Pred. No. 1.6;
 Matches 10; Conservative 6; Mismatches 6; Indels 4; Gaps 1;

QY 1 CGEGYQSRVDH-----PHFPKPIVRSI 22
 |||| :|| : || :|| :

DB 256 CGEGVKARVSVSPDEFTPPVIYSL 281
 RESULT 3
 Q9NKKD7 ID Q9NKKD7 PRELIMINARY; PRT; 648 AA.
 AC Q9NKKD7;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE HYPOTHEICAL 69.4 KDA PROTEIN.
 GN BG:DS00180.14 OR CGI8146.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Y, CN BW SP;
 RX MEDLINE=99403001; PubMed=10471707;
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
 RA Celniker S., Rubin G.M.;
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of
Drosophila melanogaster: the *Adh* region.";
 RL Genetics 153:179-219(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y, CN BW SP;
 RA Celniker S.E., Agbayani A., Arcalini T.T., Baxter E., Blazej R.G.,
 RA Butenheff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
 RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
 RA Sethi H., Snir E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,
 RA Zieran L.L., Rubin G.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003408; AAF44845.1; -;
 DR FlyBase; FBgn0028939; BG:DS00180.14.
 DR InterPro; IPR000561; EGF-like.
 DR SMART; SM00181; EGF; 10.
 DR SMART; SM00001; EGF_like; 2.
 DR PROSITE; PS01186; EGF_2; UNKNOWN_6.
 KW Hypothetical protein.
 SQ SEQUENCE 648 AA; 69419 MW; 3F16E0EBF4B94CF8 CRC64;

Query Match 37.2%; Score 54; DB 5; Length 648;
 Best Local Similarity 69.2%; Pred. No. 3.4;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPH 13
 |||| |||| |
 DB 45 CGKGYVSRKDHGH 57

RESULT 4

Q9VJU4 ID Q9VJU4 PRELIMINARY; PRT; 701 AA.
 AC Q9VJU4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CGI8146 PROTEIN.
 GN BG:DS00180.14 OR CGI8146.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX	NCBI_TaxID=7227;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=BERKELEY;
RX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA	Balwell R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Dawley S., Dahlke C., Davenport L.B., Davies P.,
RA	de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Kocha S., Dunkov B.C., Dunn P.,
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Flores C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hosdin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA	Virskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye J., Yeh R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT	"The genome sequence of Drosophila melanogaster."
RL	Science 287:2185-2195(2000).
DR	EMBL; AF003642; AAF53366.1; -
DR	FlyBase; FBgn0028939; BG:DS00180.14.
DR	InterPro; IPR000561; EGF-like.
DR	SMART; SM00181; EGF; 10.
DR	SMART; SM00001; EGF-like; 2.
DR	PROSITE; PS01186; EGF_2; UNKNOWN.6.
SQ	SEQUENCE 701 AA; 75366 MW; 95159C0AC9C09BD4 CRC64;
Query Match 37.2%; Score 54; DB 5; Length 701;	
Best Local Similarity 69.2%; Pred. No. 3.7; Indels 0; Gaps 0;	
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 CGEGYSRVDHPH 13
DB	II:II II II I CGKGYSRKDGHG 110
RESULT 5	
ID	Q30827 PRELIMINARY; PRT; 217 AA.
AC	Q30827;
DT	01-NOV-1996 (TREMBLrel. 01, Created)
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	OVAR-DOB PROTEIN (FRAGMENT).
GN	OVAR-DOB.
OS	Ovis aries (Sheep).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX	NCBI_TaxID=9940;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FINISH LANDRACE; TISSUE=BLOOD;
RX	MEDLINE=96128257; PubMed=8537127; K.T.;
RA	Wright H., Redmond J., Ballingall K.T.;
RT	"The sheep orthologue of the HLA-DQB gene.";
RL	Immunogenetics 43:76-79(1996).
DR	EMBL; 249879; CAA90036.1; -
DR	EMBL; 249880; CAA90036.1; JOINED.
DR	HSSP; P13760; 2SEB.
DR	InterPro; IPR003597; Ig-cl.
DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR000353; MHC_II_beta.
DR	Pfam; PF00047; ig; 1.
DR	Pfam; PF00969; MHC_II_beta; 1.
DR	ProbDom; PD000328; MHC_II_beta; 1.
DR	SMART; SM00407; IGcl1; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN.1.
KW	Glycoprotein; MHC II; Transmembrane.
FT	NON_TER 1
SQ	SEQUENCE 245 AA; 28069 MW; 6B21DADB2A4A299D CRC64;
Query Match 35.2%; Score 51; DB 7; Length 245;	
Best Local Similarity 47.1%; Pred. No. 3.6; Indels 0; Gaps 0;	
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;	
QY	2 GEGYSRVDPHPKPPI 18

```

Db 176 GEVYTCQVEHPSPV 192
      ||| :||| | |
      176 GEVYTCQVEHPSPV 192

RESULT 7
Q9TQA7 PRELIMINARY; PRT; 261 AA.
AC Q9TQA7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CLASS II MHC RT1.D(A) BETA CHAIN PRECURSOR.
GN RT1.D(A).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACI;
RX MEDLINE=99299366; PubMed=10369938;
RA Tian L., Wang M., Yu J., Kahan B.D., Stepkowski S.M.;
RT "Nucleotide sequences of three distinct complementary DNA clones
RT encoding rat class II major histocompatibility complex RT1.D beta-
RT chain proteins."
RL Immunogenetics 49:735-737(1999).
DR EMBL; AF084932; AAD39082.1; -.
DR HSSP; P13760; 2SEB.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Glycoprotein; MHC II; Signal; Transmembrane.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 261 CLASS II MHC RT1.D(A) BETA CHAIN.
SQ SEQUENCE 261 AA; 29496 MW; 48E220A69B374B3D CRC64;

Query Match 35.2%; Score 51; DB 7; Length 261;
Best Local Similarity 47.1%; Pred. No. 3.9;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
      ||| :||| | |
Db 192 GEVYTCQVEHPSPV 208

RESULT 8
Q9TQA5 PRELIMINARY; PRT; 264 AA.
AC Q9TQA5;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CLASS II MHC RT1.D(U) BETA CHAIN PRECURSOR.
GN RT1.D(U).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR FURTH;
RX MEDLINE=99299366; PubMed=10369938;
RA Tian L., Wang M., Yu J., Kahan B.D., Stepkowski S.M.;
RT "Nucleotide sequences of three distinct complementary DNA clones
RT encoding rat class II major histocompatibility complex RT1.D beta-
RT chain proteins."
RL Immunogenetics 49:735-737(1999).

Db 176 GEVYTCQVEHPSPV 192
      ||| :||| | |
      176 GEVYTCQVEHPSPV 192

Query Match 35.2%; Score 51; DB 7; Length 261;
Best Local Similarity 47.1%; Pred. No. 3.9;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
      ||| :||| | |
Db 192 GEVYTCQVEHPSPV 208

RESULT 9
Q9A59 PRELIMINARY; PRT; 140 AA.
AC Q9A59;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE E2 GLYCOPROTEIN (FRAGMENT).
GN E2
OS bovine viral diarrhea virus strain 4998/89.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=145184;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4998/89;
RA Greiser-Wilke I.M.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4998/89;
RA Tajima M.;
RT "Prevalence of genotypes 1 and 2 of bovine viral diarrhea virus in
RT Lower Saxony, Germany."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302959; CAC24757.1; -.
FT NON_TER 1 140
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15756 MW; 1A28D06B74391F74 CRC64;

Query Match 34.5%; Score 50; DB 12; Length 140;
Best Local Similarity 36.0%; Pred. No. 2.9;
Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 1;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
      || :| :| :| :| :|
Db 69 CGHNFRRKEDLHPY-----IGKC 87

RESULT 10
Q95536 PRELIMINARY; PRT; 142 AA.
AC Q95536;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POPY-DRB5*0603 PROTEIN (FRAGMENT).
```

```

DR EMBL; AF084934; AAD39084.1; -.
DR HSSP; P13760; 2SEB.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Glycoprotein; MHC II; Signal; Transmembrane.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 264 CLASS II MHC RT1.D(U) BETA CHAIN.
SQ SEQUENCE 264 AA; 30138 MW; 3C888533514F531F CRC64;

Query Match 35.2%; Score 51; DB 7; Length 264;
Best Local Similarity 47.1%; Pred. No. 3.9;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
      ||| :||| | |
Db 195 GEVYTCQVEHPSPV 211

RESULT 9
Q9A59 PRELIMINARY; PRT; 140 AA.
AC Q9A59;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE E2 GLYCOPROTEIN (FRAGMENT).
GN E2
OS bovine viral diarrhea virus strain 4998/89.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=145184;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=4998/89;
RA Greiser-Wilke I.M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4998/89;
RA Tajima M.;
RT "Prevalence of genotypes 1 and 2 of bovine viral diarrhea virus in
RT Lower Saxony, Germany."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302959; CAC24757.1; -.
FT NON_TER 1 140
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15756 MW; 1A28D06B74391F74 CRC64;

Query Match 34.5%; Score 50; DB 12; Length 140;
Best Local Similarity 36.0%; Pred. No. 2.9;
Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 1;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
      || :| :| :| :| :|
Db 69 CGHNFRRKEDLHPY-----IGKC 87

RESULT 10
Q95536 PRELIMINARY; PRT; 142 AA.
AC Q95536;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POPY-DRB5*0603 PROTEIN (FRAGMENT).
```

GN POPY-DRB5*0603.
 OS Pongo pygmaeus (Orangutan).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
 OX NCBI_TaxID=9600;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=931717102; PubMed=8439728;
 RA Schonbach C., Vincek V., Mayer W.E., Golubic M., O'hugin C.,
 RA Klein J.;
 RT "Multiplication of Mhc-DRB5 loci in the orangutan: implications for
 RT the evolution of DRB haplotypes";
 RL Mamm. Genome 4:159-170(1993).
 DR EMBL: S56369; AAD13881.1; -;
 DR HSSP: P13758; LAOD.
 DR InterPro: IPR003597; Ig-cl.
 DR InterPro: IPR003006; Ig_MHC.
 DR Pfam: PF00047; Ig; 1.
 DR SMART: SM00407; Igcl; 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 142 AA; 15525 MW; 02C2C3AED44D063D CRC64;

Query Match 34.5%; Score 50; DB 7; Length 142;
 Best Local Similarity 47.1%; Pred. No. 2.9;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
 ||||| :||| :|
 DB 73 GEGYTCQVHPSPVTSPL 89

RESULT 11
 Q9ESH2 PRELIMINARY; PRT; 314 AA.
 AC Q9ESH2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE PRECONDITIONING-INDUCIBLE GENE 1 PROTEIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laser M., Li Y., Xu L., Darden A., Wu B.X., Hazard E.S. III,
 RA Crosson C., Ma J.X.;
 RT "Identification and characterization of a novel gene induced by
 RT ischemic preconditioning in the retina";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF15224; AAG09182.1; -;
 SQ SEQUENCE 314 AA; 35838 MW; 4C64B70F7E909BDF CRC64;

Query Match 34.5%; Score 50; DB 11; Length 314;
 Best Local Similarity 43.8%; Pred. No. 6.8;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 10 DHPHPKPIVRSITKC 25
 :||| | :||| :|
 DB 249 NHPHPPLMCSVHPC 264

RESULT 12
 Q95M34 PRELIMINARY; PRT; 337 AA.
 AC Q95M34;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE IMMUNOGLOBULIN GAMMA 1 HEAVY CHAIN CONSTANT REGION

DE (FRAGMENT).
 GN IGHCL.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wagner B.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98383416; PubMed=9717671;
 RA Wagner B., Overesch G., Sheoran A., Holmes M., Richards C.,
 RA Leibold W., Radbruch A.;
 RT "Organization of the equine immunoglobulin heavy chain constant region
 RT genes. III. Alignment of c-mu, c-gamma, c-epsilon and c-alpha genes.";
 RL Immunobiology 199:105-119(1998).
 DR EMBL: AJ300675; CAC44624.1; -;
 FT NON_TER 1
 SQ SEQUENCE 337 AA; 37438 MW; A60BF2B01DEFD1F6 CRC64;

Query Match 34.5%; Score 50; DB 6; Length 337;
 Best Local Similarity 39.1%; Pred. No. 7.4;
 Matches 9; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
 | : : : : :
 DB 204 GKEFKCKVNNQALPQPIERTITK 226

RESULT 13
 Q947H5 PRELIMINARY; PRT; 357 AA.
 AC Q947H5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PLASMODESMAL RECEPTOR.
 GN PDR40-1.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee J.-Y., Yoo B.-C., Rojas M., Ospina N.G., Staehelin A.L.,
 RA Lucas W.J.;
 RT "Plasmodesmal receptor, NtrPD40, mediates the trafficking of movement
 RT proteins";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF307094; AAL09397.1; -;
 KW Receptor.
 SQ SEQUENCE 357 AA; 39266 MW; 295627E59FEFF3774 CRC64;

Query Match 34.5%; Score 50; DB 10; Length 357;
 Best Local Similarity 47.1%; Pred. No. 7.9;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 EGYQSRVDHPHPKPIV 19
 :| :| :| :| :|
 DB 322 QGFPDAVNHFPSTIV 338

RESULT 14
 Q947H4 PRELIMINARY; PRT; 357 AA.
 AC Q947H4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PLASMODESMAL RECEPTOR.
GN PDR40-2.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-Y., Yoo B.-C., Rojas M., Ospina N.G., Staehelin A.L.,
RA Lucas W.J.;
RT "Plasmodesmal receptor, NTRPD40, mediates the trafficking of movement
RT proteins.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF307095; AAL09398.1; -
KW Receptor.
SQ SEQUENCE 357 AA; 39276 MW; 68B7307CE83DE750 CRC64;

Query Match 34.5%; Score 50; DB 10; Length 357;
Best Local Similarity 47.1%; Pred. No. 7.9;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 EGYQSRVDHPHFPKPIV 19
: : : : :
Db 323 QGFDAVNHFPSTIV 339

RESULT 15
Q965D2
ID Q965D2 PRELIMINARY; PRT; 539 AA.
AC Q965D2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PUTATIVE CATIONIC AMINOACID TRANSPORTER.
OS Trypanosoma brucei.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RA Pereira C.A., Alonso G.D., Paveto C.M., Flavia M.M., Torres H.N.;
RT "Putative Trypanosoma brucei cationic aminoacid transporter.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF250044; AAK65959.1; -
SQ SEQUENCE 539 AA; 58047 MW; E11632AE661EC650 CRC64;

Query Match 34.5%; Score 50; DB 5; Length 539;
Best Local Similarity 42.1%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHFPKPIV 19
: : : : :
Db 425 CVSMWKIRIDHPFHFPV 443

Search completed: August 26, 2002, 10:16:19
Job time: 710 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:07:01 ; Search time 82.75 Seconds

(Without alignments)
33.557 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
1	145	100.0	25	AA1980000	Optimised IgE-CH3
2	122	84.1	313	AA199996	Rat immunoglobulin
3	122	84.1	340	AA199996	Rat IgE heavy chain
4	122	84.1	341	AA199996	Immunogenic peptide
5	121	83.4	45	AA199996	IgE immunogenic pe
6	121	83.4	46	AA199996	IgE immunogenic pe
7	106	73.1	25	AA199996	Optimised IgE-CH3
8	105	72.4	25	AA199996	IgE immunogenic pe
9	105	72.4	45	AA199996	IgE immunogenic pe
10	105	72.4	45	AA199996	IgE immunogenic pe
11	105	72.4	46	AA199996	IgE immunogenic pe

12	105	72.4	57	21	AA1980081	IgE immunogenic pe
13	105	72.4	62	21	AA1980080	IgE immunogenic pe
14	105	72.4	63	21	AA1980084	IgE immunogenic pe
15	103	71.0	25	21	AA199998	Optimised IgE-CH3
16	103	71.0	25	21	AA191212	Modified human IgE
17	103	71.0	25	21	AA1980012	Peptide sequence o
18	103	71.0	42	21	AA1980014	IgE immunogenic pe
19	103	71.0	42	21	AA191215	Modified MVF Th ep
20	103	71.0	42	21	AA191216	Modified MVF Th ep
21	103	71.0	42	21	AA191217	Modified MVF Th ep
22	103	71.0	42	21	AA191218	Modified MVF Th ep
23	103	71.0	45	21	AA1980007	IgE-CH3 domain ant
24	103	71.0	45	21	AA191218	Modified HBV surfa
25	103	71.0	45	21	AA1986005	Peptide sequence o
26	103	71.0	46	21	AA1980011	IgE immunogenic pe
27	103	71.0	46	21	AA191213	Modified MVF Th ep
28	103	71.0	46	21	AA191214	Modified MVF Th ep
29	103	71.0	46	21	AA1986003	Peptide sequence o
30	103	71.0	56	21	AA1980016	IgE immunogenic pe
31	103	71.0	59	21	AA1980010	IgE immunogenic pe
32	103	71.0	60	21	AA1980013	IgE immunogenic pe
33	103	71.0	60	21	AA1980015	IgE immunogenic pe
34	103	71.0	63	21	AA1980008	IgE-CH3 domain ant
35	103	71.0	63	21	AA1980012	IgE immunogenic pe
36	103	71.0	63	21	AA191219	Inv epitope/modifi
37	103	71.0	63	21	AA1986006	Peptide sequence o
38	96	66.2	60	21	AA1980078	IgE immunogenic pe
39	90	62.1	313	21	AA199997	Mouse immunoglobul
40	89	61.4	25	21	AA1980077	Optimised IgE-CH3
41	85	58.6	561	20	AA1917415	Mouse immunoglobul
42	84	57.9	110	14	AA1933317	Variant IgE - muta
43	82	56.6	124	18	AA1924097	Partial canine imm
44	82	56.6	312	21	AA199995	Dog immunoglobulin
45	82	56.6	341	21	AA19806208	Immunogenic peptid

ALIGNMENTS

RESULT 1

AA1980000
ID AA1980000 standard; Peptide; 25 AA.

XX
XX AA1980000;

XX
DT 15-MAY-2000 (first entry)

XX
DE Optimised IgE-CH3 domain antigen peptide for rat IgE.

XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX
OS Rattus sp.

XX
OS Synthetic.

XX
PN WO967293-A1.

XX
PD 29-DEC-1999.

XX
PF 21-JUN-1999; 99WO-US13959.

XX
PR 20-JUN-1998; 98US-0100287.

XX
PA (UNBI-) UNITED BIOMEDICAL INC.

XX
PI Wang CY, Walfield AM;

XX
DR WPI; 2000-160578/14.

XX
PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -

XX PS Claim 1; Page 99; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 145; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 7.2e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHFKPIVRSITKC 25

Db 1 cgegysrvdhpfpkplvrsitkc 25

RESULT 2

AAY79996

ID AAY79996 standard; Protein; 313 AA.

XX AC AAY79996;

XX DT 15-MAY-2000 (first entry)

XX DE Rat immunoglobulin E epsilon heavy chain SEQ ID NO:3.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Rattus sp.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy -

XX PS Example 1; Page 66-68; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 313 AA;

Query Match 84.1%; Score 122; DB 21; Length 313;

Best Local Similarity 95.7%; Pred. No. 3.3e-10;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHFKPIVRSITK 24

Db 190 gegyqsrvdhpfpkplvrsitk 212

RESULT 3

AAB03643

ID AAB03643 standard; protein; 340 AA.

XX AC AAB03643;

XX DT 22-NOV-2000 (first entry)

XX DE Rat IgE heavy chain constant regions 2, 3 and 4.

XX KW Rat; immunoglobulin E; IgE; vaccination; infection; allergy;

XX KW asthma; eczema; immunogenic peptide.

XX OS Rattus sp.

XX PN WO200025722-A2.

XX PD 11-MAY-2000.

XX PF 21-OCT-1999; 99WO-SE01896.

XX PR 02-NOV-1998; 98US-0106552.

XX PR 22-SEP-1999; 99US-0401636.

XX PA (RESI-) RESISTENTIA PHARM AB.

XX PI Hellman LT;

XX DR WPI; 2000-365342/31.

XX PT Immunogenic polypeptides useful for preventing the harmful effects of

XX PT immunoglobulin E in mammals -

XX PS Disclosure; Fig 1; 50pp; English.

XX CC The present sequence is an immunogenic peptide consisting of the

XX CC heavy chain constant regions 2, 3 and 4 of the rat IgE. It was used to

XX CC construct a number of immunogenic peptides which consisted of regions of

XX CC IgE from different mammals, which appear to cause a stronger polyclonal

XX CC anti-self IgE response than peptides consisting of the same regions from

XX CC one mammal. Immunogenic peptides, particularly those consisting of

XX CC different heavy chain constant regions, can be used for vaccination

XX CC in humans, against bacterial and viral infections and allergies, such

XX CC as asthma, fur, pollen and food allergies and eczema.

XX SQ Sequence 340 AA;

Query Match 84.1%; Score 122; DB 21; Length 340;
Best Local Similarity 95.7%; Pred. No. 3.6e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
||||| ||||||| ||||||| |||||||
Db 195 gegyqsrvdhphfhpkipvrsitk 217

RESULT 4

AAB06206
ID AAB06206 standard; protein; 341 AA.

AC AAB06206;

XX 22-NOV-2000 (first entry)

DE Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.

KW Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW asthma; eczema; immunogenic peptide.

XX Chimeric - Didelphis virginiana.

OS Chimeric - Rattus sp.

XX WO200025722-A2.

PN 11-MAY-2000.

XX 21-OCT-1999; 99WO-S001896.

PR 02-NOV-1998; 98US-0106652.

PR 22-SEP-1999; 99US-0401636.

XX (RESI-) RESISTENTIA PHARM AB.

XX Hellman LT;

XX WPI; 2000-365342/31.

XX Immunogenic polypeptides useful for preventing the harmful effects of
PT immunoglobulin E in mammals -

PS Disclosure; Fig 2; 50pp; English.

XX The present sequence is an immunogenic peptide consisting of the heavy
CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC constant region 3 from the rat. It was shown to cause a stronger
CC polyclonal anti-self IgE response than peptides consisting of the same
CC regions from one mammal. Immunogenic peptides, particularly those
CC consisting of different heavy chain constant regions, can be used for
CC vaccination in humans, against bacterial and viral infections and
CC allergies, such as asthma, fur, pollen and food allergies and eczema.

XX Sequence 341 AA;

Query Match 84.1%; Score 122; DB 21; Length 341;
Best Local Similarity 95.7%; Pred. No. 3.6e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
||||| ||||||| ||||||| |||||||
Db 205 gegyqsrvdhphfhpkipvrsitk 228

RESULT 5

AAY80018
ID AAY80018 standard; Peptide; 45 AA.

XX AAY80018;

XX antibody; allergy; allergic disease; immunisation; anti-allergic;

DT 15-MAY-2000 (first entry)
XX IgE immunogenic peptide conjugate SEQ ID NO:25.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -

XX Claim 14; Page 76; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 45 AA;

Query Match 83.4%; Score 121; DB 21; Length 45;
Best Local Similarity 88.0%; Pred. No. 5.5e-11;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITK 25

Db 21 cgygyqsrvdhphfhpkipvrsitkc 45

RESULT 6

AAY80017

ID AAY80017 standard; Peptide; 46 AA.

XX AAY80017;

XX 15-MAY-2000 (first entry)

DE IgE immunogenic peptide conjugate SEQ ID NO:24.

KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 OS Unidentified.
 XX
 XX
 PN WO9967293-A1.
 XX
 XX 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 DR
 XX
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 for immunization against allergy -
 XX
 PS Claim 14; Page 76; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 and anti-asthmatic properties. (I) induces polyclonal antibodies
 specific for a target effector site on the epsilon-heavy chain of IgE,
 and so preventing triggering and activation of mast cells and basophils
 and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 containing (I) are used for active immunisation against IgE-mediated
 allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 dermatitis. Nucleic acids that encode these compounds are useful for
 recombinant production of corresponding peptides or in DNA vaccines.
 Conjugates of (I) that include a promiscuous T helper cell epitope
 (functional in genetically diverse subjects), in addition to a B cell
 target epitope, have increased immunogenicity and may include cyclic
 constraints (disulfide bridge) to stabilise conformational features and
 maximize cross-reactivity to the natural target. They induce safe
 (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 acid sequences used in the exemplification of the present invention.
 XX Sequence 46 AA;
 SQ

Query Match 83.4%; Score 121; DB 21; Length 46;
 Best Local Similarity 88.0%; Pred. No. 5.7e-11;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHFKPIVRSITKC 25
 II IIII IIII IIIIIIIIIII
 Db 22 cgygysivdhpdpfpkpvrsitkc 46

RESULT 7
 AAY80001
 ID AAY80001 standard; Peptide; 25 AA.
 XX
 AC AAY80001;
 XX
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.
 XX
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Mus sp.
 OS Synthetic.
 XX
 PN WO9967293-A1.
 XX

PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 DR
 XX
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 for immunization against allergy -
 XX
 PS Claim 1; Page 100; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 and anti-asthmatic properties. (I) induces polyclonal antibodies
 specific for a target effector site on the epsilon-heavy chain of IgE,
 and so preventing triggering and activation of mast cells and basophils
 and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 containing (I) are used for active immunisation against IgE-mediated
 allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 dermatitis. Nucleic acids that encode these compounds are useful for
 recombinant production of corresponding peptides or in DNA vaccines.
 Conjugates of (I) that include a promiscuous T helper cell epitope
 (functional in genetically diverse subjects), in addition to a B cell
 target epitope, have increased immunogenicity and may include cyclic
 constraints (disulfide bridge) to stabilise conformational features and
 maximize cross-reactivity to the natural target. They induce safe
 (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 acid sequences used in the exemplification of the present invention.
 XX Sequence 25 AA;
 SQ

Query Match 73.1%; Score 106; DB 21; Length 25;
 Best Local Similarity 80.0%; Pred. No. 5.2e-09;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHFKPIVRSITKC 25
 II IIII IIII IIIIIIIIIII
 Db 1 cgygysivdhpdpfpkpvrsitkc 25

RESULT 8
 AAY79999
 ID AAY79999 standard; Peptide; 25 AA.
 XX
 AC AAY79999;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for dog IgE.
 XX
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Canis sp.
 OS Synthetic.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX

PA (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY, Walfield AM;
 PI WPI; 2000-160578/14.
 DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 XX for immunization against allergy -
 PT Claim 1; Page 99; 155pp; English.
 PS The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX SQ Sequence 25 AA;
 Query Match 72.4%; Score 105; DB 21; Length 25;
 Best Local Similarity 76.0%; Pred. No. 7.3e-09; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 6;
 QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 III I III III I III III I
 Db 1 cgetysrvrthphpkdivrsiack 25
 RESULT 9
 AAY80019
 ID AAY80019 standard; Peptide; 45 AA.
 AC AAY80019;
 XX 15-MAY-2000 (first entry)
 DT IgE immunogenic peptide conjugate SEQ ID NO:26.
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX Unidentified.
 OS WO9967293-A1.
 XX 29-DEC-1999.
 XX 21-JUN-1999; 99WO-US13959.
 XX 20-JUN-1998; 98US-0100287.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA Wang CY, Walfield AM;
 PI WPI; 2000-160578/14.
 DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 XX for immunization against allergy -
 PT Claim 1; Page 99; 155pp; English.

PT for immunization against allergy -
 XX Claim 14; Page 76; 155pp; English.
 PS The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX SQ Sequence 45 AA;
 Query Match 72.4%; Score 105; DB 21; Length 45;
 Best Local Similarity 76.0%; Pred. No. 1.4e-08;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 III I III III I III III I
 Db 21 cgetysrvrthphpkdivrsiack 45
 RESULT 10
 AAY80083
 ID AAY80083 standard; Peptide; 45 AA.
 AC AAY80083;
 XX 15-MAY-2000 (first entry)
 DT IgE immunogenic peptide conjugate SEQ ID NO:90.
 DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX Unidentified.
 OS WO9967293-A1.
 XX 29-DEC-1999.
 XX 21-JUN-1999; 99WO-US13959.
 XX 20-JUN-1998; 98US-0100287.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA Wang CY, Walfield AM;
 PI WPI; 2000-160578/14.
 DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 XX for immunization against allergy -
 PT Claim 14; Page 77; 155pp; English.
 PS The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 45 AA;

Query Match 72.4%; Score 105; DB 21; Length 45;
 Best Local Similarity 76.0%; Pred. No. 1.4e-08;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITK 25
 III I III III I IIIII I I
 Db 21 cgetysrvthphlpkdivrskiakc 45

RESULT 11
 AAY80020
 ID AAY80020 standard; Peptide; 46 AA.

XX
 AC AAY80020;
 DT 15-MAY-2000 (first entry)
 XX

DE IgE immunogenic peptide conjugate SEQ ID NO:27.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX

PS Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 46 AA;

Query Match 72.4%; Score 105; DB 21; Length 46;
 Best Local Similarity 76.0%; Pred. No. 1.4e-08;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITK 25
 III I III III I IIIII I I
 Db 22 cgetysrvthphlpkdivrskiakc 46

RESULT 12
 AAY80081
 ID AAY80081 standard; Peptide; 57 AA.

XX
 AC AAY80081;
 XX 15-MAY-2000 (first entry)
 DT
 XX

DE IgE immunogenic peptide conjugate SEQ ID NO:88.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX

PS Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

RESULT 15

AA79998
ID AAY79998 standard; Peptide; 25 AA.
XX AC AAY79998;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for human IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Homo sapiens.
OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
XX PS Claim 1; Page 21; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.5e-08;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPRKPIVRSTKTC 25
||| ||||| ||| | :|| |||
Db 1 cgetygsrvthphlpralmrsttkc 25

Search completed: August 26, 2002, 10:07:02
Job time: 228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:05:18 ; Search time 34.18 seconds
(without alignments)
17.865 Million cell updates/sec

Title: US-09-701-623c-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPKPIVRSTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	103	71.0	25	3	US-09-100-414B-95
2	103	71.0	25	4	US-09-303-323-95
3	103	71.0	42	3	US-09-100-414B-98
4	103	71.0	42	3	US-09-100-414B-99
5	103	71.0	42	3	US-09-100-414B-100
6	103	71.0	42	4	US-09-303-323-98
7	103	71.0	42	4	US-09-303-323-99
8	103	71.0	42	4	US-09-303-323-100
9	103	71.0	45	3	US-09-100-414B-101
10	103	71.0	45	4	US-09-303-323-101
11	103	71.0	46	3	US-09-100-414B-96
12	103	71.0	46	3	US-09-100-414B-97
13	103	71.0	46	4	US-09-303-323-96
14	103	71.0	46	4	US-09-303-323-97
15	103	71.0	63	3	US-09-100-414B-102
16	103	71.0	63	4	US-09-303-323-102
17	85	58.6	561	3	US-09-192-545-2
18	82	56.6	426	1	US-08-336-583-2
19	82	56.6	426	5	PCT-US95-13795-2
20	80	55.2	106	2	US-08-232-539D-54
21	80	55.2	113	2	US-08-232-539D-56
22	69.5	47.9	119	2	US-08-464-025A-1
23	64	44.1	109	4	US-08-466-163B-1
24	62	42.8	22	2	US-08-232-539D-19
25	62	42.8	24	2	US-08-232-539D-20
26	62	42.8	56	2	US-08-232-539D-18
27	53.5	36.9	118	3	US-08-466-151-1

28 53 36.6 331 2 US-08-646-981-17 Sequence 17, Appl
29 53 36.6 447 6 5455030-1 Patent No. 5455030
30 49 33.8 338 3 US-08-890-719-12 Sequence 12, Appl
31 49 33.8 355 3 US-08-890-719-11 Sequence 11, Appl
32 49 33.8 355 3 US-08-890-719-13 Sequence 13, Appl
33 48 33.1 212 4 US-08-811-463-39 Sequence 39, Appl
34 48 33.1 320 2 US-08-579-940-8 Sequence 8, Appl
35 48 33.1 324 2 US-08-579-940-7 Sequence 7, Appl
36 48 33.1 445 1 US-08-353-400-33 Sequence 33, Appl
37 48 33.1 464 1 US-08-353-400-36 Sequence 36, Appl
38 48 33.1 599 1 US-08-442-542-18 Sequence 18, Appl
39 48 33.1 599 3 US-08-765-469-18 Sequence 18, Appl
40 46 31.7 373 2 US-08-864-804-3 Sequence 3, Appl
41 45 31.0 357 3 US-08-630-172-20 Sequence 20, Appl
42 45 31.0 357 4 US-09-375-419-20 Sequence 20, Appl
43 45 31.0 367 3 US-08-630-172-19 Sequence 19, Appl
44 45 31.0 367 4 US-09-375-419-19 Sequence 19, Appl
45 45 31.0 378 3 US-08-630-172-21 Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 71.0%; Score 103; DB 3; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSTKC 25
||| ||||| ||| | : || ||||
Db 1 CGETYQSRVTHPLPALMRSTKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 622897
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 71.0%; Score 103; DB 4; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHFKPKIVRSITKC 25
||| ||||| ||| | :|| |||
Db 1 CGETYQSRVTHPLPALMRSTKC 25

RESULT 3
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHFKPKIVRSITKC 25
||| ||||| ||| | :|| |||
Db 18 CGETYQSRVTHPLPALMRSTKC 42

RESULT 4
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 3e-09;

Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Db 18 CGETYQSRVTHPLRALMRSTTKC 42

RESULT 5

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Db 18 CGETYQSRVTHPLRALMRSTTKC 42

RESULT 6

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 71.0%; Score 103; DB 4; Length 42;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Db 18 CGETYQSRVTHPLRALMRSTTKC 42

RESULT 7

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 71.0%; Score 103; DB 4; Length 42;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8

US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 71.0%; Score 103; DB 4; Length 42;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| |||| | : || |||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 9

US-09-100-414B-101
; Sequence 101, Application US/09100414B

; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-101

Query Match 71.0%; Score 103; DB 3; Length 45;
Best Local Similarity 68.0%; Pred. No. 3.2e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 10

US-09-303-323-101
; Sequence 101, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 71.0%; Score 103; DB 4; Length 45;
Best Local Similarity 68.0%; Pred. No. 3.2e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | :|| |||
Db 21 CGEYOSRVTHPLPALMRSTKC 45

RESULT 11
US-09-100-414B-96
Sequence 96, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-96

Query Match 71.0%; Score 103; DB 3; Length 46;
Best Local Similarity 68.0%; Pred. No. 3.3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYSRVDHPHPKPIVRSITKC 25

Db 22 CGEYOSRVTHPLPALMRSTKC 46
||| ||||| ||| | :|| |||
RESULT 12
US-09-100-414B-97
Sequence 97, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 71.0%; Score 103; DB 3; Length 46;
Best Local Similarity 68.0%; Pred. No. 3.3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | :|| |||
Db 22 CGEYOSRVTHPLPALMRSTKC 46

RESULT 13
US-09-303-323-96
Sequence 96, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Y1
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

;
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-96

Query Match 71.0%; Score 103; DB 4; Length 46;
Best Local Similarity 68.0%; Pred. No. 3.3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGEGYOSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 22 CGETYOSRVTHPLPALMRSTTKC 46

RESULT 14
US-09-303-323-97
; Sequence 97, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid

;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-97

Query Match 71.0%; Score 103; DB 4; Length 46;
Best Local Similarity 68.0%; Pred. No. 3.3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGEGYOSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 22 CGETYOSRVTHPLPALMRSTTKC 46

RESULT 15
US-09-100-414B-102
; Sequence 102, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-102

Query Match 71.0%; Score 103; DB 3; Length 63;
Best Local Similarity 68.0%; Pred. No. 4.6e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGEGYOSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 39 CGETYOSRVTHPLPALMRSTTKC 63

Search completed: August 26, 2002, 10:05:18
Job time: 124 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:04:25 ; Search time 45.87 Seconds
(without alignments)
52.370 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYGQSIVD RDPFKPIVRSITLC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database :      PIR_71:**
1:  pir1:**
2:  pir2:**
3:  pir3:**
4:  pir4:**
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	116	82.9	423	1	EHMS5	Ig epsilon chain C
2	108	77.1	107	2	I68730	IgE chain C3 regio
3	108	77.1	107	2	I68726	IgE chain C3 regio
4	108	77.1	388	1	EHMS	Ig epsilon chain C
5	108	77.1	548	2	S38864	Ig epsilon chain C
6	85	60.7	429	1	EHRT	Ig epsilon chain C
7	49	35.0	165	2	T31043	hypothetical prote
8	49	35.0	232	2	AQ2525	hypothetical prote
9	49	35.0	433	2	C75354	conserved hypothe
10	49	35.0	705	2	A75026	phosphoribosylform
11	48.5	34.6	363	1	BVBK2	MAK32 protein - ye
12	48	34.3	169	1	I12142	cytochrome-c oxida
13	48	34.3	294	2	G72293	phosphate butyrylt
14	48	34.3	1259	2	T06521	pitriylsin (EC 3:4
15	47	33.6	169	1	A35209	cytochrome-c oxida
16	47	33.6	241	2	E71532	probable porphobil
17	47	33.6	591	2	B59563	probable acetolact
18	47	33.6	914	2	A00311	NADH dehydrogenase
19	46.5	33.2	210	2	C75399	hypothetical prote
20	46	32.9	172	2	A23378	adenine phosphorib
21	46	32.9	351	2	T08001	probable fadE36 pr
22	46	32.9	359	2	F89075	protein K04A8.1 [i
23	46	32.9	426	2	I36948	Ig epsilon chain -
24	46	32.9	430	2	AG0531	cell cycle protein
25	46	32.9	705	2	A71211	probable phosphori
26	46	32.9	715	2	S54628	hypothetical prote
27	45	32.1	172	2	S75440	adenine phosphorib
28	45	32.1	294	2	S72706	probable enoyl-CoA
29	45	32.1	910	2	G91024	NADH dehydrogenase

30	45	32.1	910	2	A65000	NADH dehydrogenase
31	45	32.1	910	2	H85868	NADH dehydrogenase
32	45	32.1	910	2	A10796	NADH dehydrogenase
33	45	32.1	921	2	G02326	transcription fact
34	44.5	31.8	62	1	SBHUP	statherin precursor
35	44.5	31.8	222	2	S38779	beta-amylase (BC 3
36	44.5	31.8	330	1	A87198	phosphoesterase-re
37	44.5	31.8	335	2	T16904	hypothetical prote
38	44.5	31.8	1056	2	S55151	probable membrane
39	44	31.4	46	2	D49281	pol protein - simi
40	44	31.4	169	1	OLB04	cytochrome-c oxida
41	44	31.4	282	2	T39767	translation initia
42	44	31.4	587	1	KSVA0	L-ascorbate oxidas
43	44	31.4	699	2	B72775	probable DNA repli
44	44	31.4	706	2	D83112	elongation factor
45	44	31.4	1058	2	T30580	p-type ATPase - sl

ALIGNMENTS

```

RESULT 1
EHMSS
Ig epsilon chain C region (version 2) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
C:Accession: A02145
R: Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.
EMBO J. 1, 1117-1123, 1982
A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison
A:Reference number: A90966; MUID:84236092
A:Accession: A02145
A:Molecule type: DNA
A:Residues: 1-423 <ISH>
A:Note: the sequence was determined from the germline gene
C:Genetics:
A:Introns: 91/1; 199/1; 307/1
A:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:16-77/Domain: immunoglobulin homology <IMM1>
F:115-183/Domain: immunoglobulin homology <IMM2>
F:220-288/Domain: immunoglobulin homology <IMM3>
F:325-396/Domain: immunoglobulin homology <IMM4>
F:23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.9%; Score 116; DB 1; Length 423;
Best Local Similarity 91.3%; Pred. No. 7e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGYSIVDRDPFKPIVRSITL 24
|||||:|||||
Db 281 GYGYSIVDRDPFKPIVRSITL 303

RESULT 2
I68730
IgE chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: I68730
R: Shinkai, Y.; Nakachi, H.; Honjo, T.; Okumura, K.
Immunogenetics 27, 288-292, 1988
A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic aci
A:Reference number: I54443; MUID:88152907
A:Accession: I68730
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-107 <RES>
A:Cross-references: GB:M22933; NID:q194464; PIDN:AAA37915.1; PID:q194469

```

C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 77.1%; Score 108; DB 2; Length 107;
Best Local Similarity 90.9%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23

Db 83 GYGQCIVDHPDPFKPIVRSIT 104

RESULT 3

I68726

IgE chain C3 region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000

C:Accession: I68726

R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid s

A:Reference number: I54443; MUID:88152907

A:Accession: I68726

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-107 <RES>

A:Cross-references: GB:M22930; NID:g194455; PIDN:AAA37911.1; PID:g194460

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 77.1%; Score 108; DB 2; Length 107;
Best Local Similarity 90.9%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23

Db 83 GYGQCIVDHPDPFKPIVRSIT 104

RESULT 4

EHMS

Ig epsilon chain C region (version 1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999

C:Accession: A02144

R:Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.

Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982

A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.

A:Reference number: A02144; MUID:83117774

A:Accession: A02144

A:Molecule type: mRNA

A:Residues: 1-388 <LIU>

A:Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (ka

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:1-44/Domain: immunoglobulin homology (fragment) <IM1>

F:81-149/Domain: immunoglobulin homology <IM2>

F:186-254/Domain: immunoglobulin homology <IM3>

F:290-361/Domain: immunoglobulin homology <IM4>

F:10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 77.1%; Score 108; DB 1; Length 388;
Best Local Similarity 90.9%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23

Db 247 GYGQCIVDHPDPFKPIVRSIT 268

RESULT 5

S38864

Ig epsilon chain C region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001

C:Accession: S38864

R:Kipp, B.; Becker, W.; Schlaak, M.

submitted to the EMBL Data Library, November 1993

A:Description: Combination of a defined specificity and desired isotype by cloning of

A:Reference number: S38864

A:Accession: S38864

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-548 <KIP>

A:Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:353-421/Domain: immunoglobulin homology <IMM>

Query Match 77.1%; Score 108; DB 2; Length 548;
Best Local Similarity 90.9%; Pred. No. 1.6e-08;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23

Db 414 GYGQCIVDHPDPFKPIVRSIT 435

RESULT 6

EHRT

Ig epsilon chain C region - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999

C:Accession: A93442; A90937; A02143

R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.

Nucleic Acids Res. 10, 6041-6049, 1982

A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.

A:Reference number: A93442; MUID:83064537

A:Accession: A93442

A:Molecule type: mRNA

A:Residues: 1-429 <HEL>

A:Experimental source: strain LOU/c/Wsl, Immunocytooma IR2

R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.

DNA 1, 335-343, 1982

A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction

A:Reference number: A90937; MUID:83182019

A:Contents: myeloma IRI62

A:Accession: A90937

A:Molecule type: mRNA

A:Residues: 'N', 169-307, 'L', 309-342 <KIN>

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into (

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:19-80/Domain: immunoglobulin homology <IM1>

F:118-186/Domain: immunoglobulin homology <IM2>

F:223-291/Domain: immunoglobulin homology <IM3>

F:327-398/Domain: immunoglobulin homology <IM4>

F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 60.7%; Score 85; DB 1; Length 429;
Best Local Similarity 77.3%; Pred. No. 4.2e-05;
Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGQSIIVDRPDPFKPIVRSIT 23

Db 284 GEGYQCRVDHPDPFKPIVRSIT 305

RESULT 7

T31043
hypothetical protein 165 - Sulfolobus sp. plasmid pNOB8
C:Species: Sulfolobus sp.
C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T31043
R:She, O.; Phan, H.; Garrett, R.A.; Albers, S.V.; Stedman, K.M.; Zillig, W.
Extremophiles 2, 417-425, 1998
A:Title: Genetic profile of pNOB8 from Sulfolobus: the first conjugative plasmid from an
A:Reference number: Z20959; MUID:99044580
A:Accession: T31043
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-165 <SHE>
A:Cross-references: EMBL:AJ010405; NID:el351926; PID:el351965; PIDN:CAA09149.1
A:Experimental source: strain NOB8H2
C:Genetics:
A:Genome: plasmid pNOB8

Query Match 35.0%; Score 49; DB 2; Length 165;
Best Local Similarity 58.8%; Pred. No. 4.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 5 YQSIDVRDPFKPIVR 21
I: I I I I I I I I I I
Db 71 YERIDKDPNPRIVRS 87
RESULT 8
AG2525
hypothetical protein alr7383 [imported] - Anabaena sp. (strain PCC 7120) plasmid pCC7120
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AG2525
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriquchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <KUR>
A:Cross-references: GB:BA000020; PIDN:BA77141.1; PID:gl7134582; GSPDB:GN00180
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr7383
A:Genome: plasmid

Query Match 35.0%; Score 49; DB 2; Length 232;
Best Local Similarity 43.8%; Pred. No. 7.2;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 5 YQSIDVRDPFKPIVR 20
I: : : : : : : : : :
Db 49 YAAVDHPDPNCVK 64
RESULT 9
C75354
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: C75354
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vanathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: C75354

A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <WHI>
A:Cross-references: GB:AE002019; NID:g6459547; PIDN:AAF11328.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1773
A:Map position: 1

Query Match 35.0%; Score 49; DB 2; Length 433;
Best Local Similarity 35.3%; Pred. No. 14;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 6 QSIDVRDPFKPIVRSI 22
:: I I I I I I I I I I
Db 117 EAVAQRDLPEPLRLQ 133

RESULT 10
A75026
phosphoribosylformylglycinamide synthase II (purl) PAB1201 - Pyrococcus abyssi (str
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: A75026
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A:Reference number: A75001
A:Accession: A75026
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-705 <KAW>
A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50647.1; PID:g545
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1201
C:Superfamily: phosphoribosylformylglycinamide synthase component II

Query Match 35.0%; Score 49; DB 2; Length 705;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 5 YQSIDVRDPFKPIVRSI 22
I: : : : : I I I I I
Db 513 YNEVDPRVKPTEPVAGI 530

RESULT 11
BVBK2
MAK32 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YCR019w
C:Species: Saccharomyces cerevisiae
C>Date: 30-Jun-1991 #sequence_revision 30-Sep-1993 #text_change 16-Jun-2000
C:Accession: S19429; S07695
R:Feldmann, H.; Mannhaupt, G.; Vetter, I.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19429
A:Accession: S19429
A:Molecule type: DNA
A:Residues: 1-363 <FEL>
A:Cross-references: EMBL:X59720; NID:gl907116; PIDN:CAA42310.1; PID:gl907162; GSPDB:G
R:Toh-e, A.; Sahashi, Y.
Yeast 1, 159-171, 1985
A:Title: The PET18 locus of Saccharomyces cerevisiae: a complex locus containing mult
A:Reference number: S07692; MUID:89131254
A:Accession: S07695
A:Molecule type: DNA
A:Residues: 1-14, 'I', 15-81, 83-282, 'S', 284-363 <TOH>
C:Genetics:
A:Gene: SGD:MAK32; MIPS:YCR019w
A:Cross-references: SGD:S0000612; MIPS:YCR019w

A:Map position: 3R
C:Superfamily: MAK32 protein

Query Match 34.6%; Score 48.5; DB 1; Length 363;
Best Local Similarity 55.0%; Pred. No. 14; Mismatches 3; Conservative 11; Gaps 1;

QY 4 YQSIIVDRPDPFKPIVRSI 22
I : |||| |||| : : I
Db 61 GLKWIIVDRGSDFPKEVIREI 80

RESULT 12
S12142
cytochrome-c oxidase (EC 1.9.3.1) chain IV precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S12142; S16114
R:Grossman, L.I.; Akamatsu, M.
Nucleic Acids Res. 18, 6454, 1990
A:Title: Nucleotide sequence of a mouse cDNA for subunit IV of cytochrome c oxidase.
A:Reference number: S12142; MUID:91057158
A:Accession: S12142
A:Molecule type: mRNA
A:Residues: 1-169 <GRO>
A:Cross-references: EMBL:X54691; NID:g50518; PIDN:CAA38507.1; PID:g50519
A:Experimental source: strain Balb/c
R:Carter, R.S.; Avadhani, N.G.
Arch. Biochem. Biophys. 288, 97-106, 1991
A:Title: Cloning and characterization of the mouse cytochrome c oxidase subunit IV gene.
A:Reference number: S16114; MUID:91378465
A:Accession: S16114
A:Molecule type: mRNA
A:Residues: 1-169 <ARC>
C:Genetics:
A:Gene: COXIV
A:Genome: nuclear
C:Superfamily: cytochrome-c oxidase chain IV
C:Keywords: membrane-associated complex; mitochondrion; oxidative phosphorylation; oxid
F:1-22/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F:23-169/Product: cytochrome-c oxidase chain IV #status predicted <MAT>

Query Match 34.3%; Score 48; DB 1; Length 169;
Best Local Similarity 36.4%; Pred. No. 7.2;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 YGQSIIVDRPDPFKPIVRSITL 24
I : || || || : : I
Db 33 YAPPTVADRRDYLPLDVAHVTM 54

RESULT 13
G72293
phosphate butyryltransferase - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72293
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316
A:Accession: G72293
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294 <ARN>
A:Cross-references: GB:AE001770; GB:AE000512; NID:g4981658; PIDN:AAD36206.1; PID:g498167
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1130

C:Superfamily: phosphate acetyltransferase

Query Match 34.3%; Score 48; DB 2; Length 294;
Best Local Similarity 46.7%; Pred. No. 13;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21
I : : : : |||| : : I
Db 122 SYWEIPDFRPLIIS 136

RESULT 14
T06521
pitriylisin (EC 3.4.24.55) - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T06521
R:Vanderveere, P.S.; Bennett, T.M.; Oblong, J.E.; Lamppa, G.K.
Proc. Natl. Acad. Sci. U.S.A. 92, 7177-7181, 1995
A:Title: A chloroplast processing enzyme involved in precursor maturation shares a zi
A:Reference number: Z15733; MUID:95365331
A:Accession: T06521
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1259 <VAN>
A:Cross-references: EMBL:U25111; NID:g1065907; PIDN:AAA81472.1; PID:g799369
C:Genetics:
A:Genome: nuclear
C:Function:
A:Description: catalyzes proteolytic removal of chloroplast transit peptides
C:Keywords: hydrolase; metalloproteinase; zinc

Query Match 34.3%; Score 48; DB 2; Length 1259;
Best Local Similarity 38.1%; Pred. No. 68;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 5 YOSIVDRPDPFKPIVRSITLC 25
I : : : || || : : I
Db 79 YSSVLSPQTPAPVPRQSCISC 99

RESULT 15
A35209
cytochrome-c oxidase (EC 1.9.3.1) chain IV precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 11-Jun-1999
C:Accession: A35209; S12724; S04070; S04593; S14190; S65373
R:Yamada, M.; Amuro, N.; Goto, Y.; Okazaki, T.
J. Biol. Chem. 265, 7687-7692, 1990
A:Title: Structural organization of the rat cytochrome c oxidase subunit IV gene.
A:Reference number: A35209; MUID:90237079
A:Accession: A35209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <VAM>
A:Cross-references: GB:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517
R:Amuro, N.; Yamada, M.; Goto, Y.; Okazaki, T.
Nucleic Acids Res. 18, 3992, 1990
A:Title: Complete nucleotide sequence of the gene encoding rat cytochrome c oxidase s
A:Reference number: S12724; MUID:90326528
A:Accession: S12724
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <AMU>
A:Cross-references: EMBL:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517
R:Goto, Y.; Amuro, N.; Okazaki, T.
Nucleic Acids Res. 17, 2851, 1989
A:Title: Nucleotide sequence of cDNA for rat brain and liver cytochrome c oxidase sub
A:Reference number: S04070; MUID:89240039
A:Accession: S04070
A:Molecule type: mRNA

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:17:03 ; Search time 24.08 Seconds
(without alignments)
40.199 Million cell updates/sec

Title: US-09-701-623c-8
Perfect score: 140
Sequence: 1 CGYGYQSIIVDRDPFKPIVRSITLC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	108	77.1	421	1 EPC_MOUSE	P06336 mus musculus
2	85	60.7	429	1 EPC_RAT	P01855 rattus norv
3	49	35.0	705	1 PURL_PYRAB	Q9uxw6 pyrococcus
4	48.5	34.5	363	1 MK32_YEAST	P23060 saccharomyc
5	48	34.3	144	1 COX4_PITPI	O46585 plthecia pi
6	48	34.3	169	1 COX4_MOUSE	P19783 mus musculu
7	48	34.3	294	1 PTA_THENA	Q9x014 thermotoga
8	47	33.6	169	1 COX4_RAT	P10888 rattus norv
9	47	33.6	241	1 HEM3_CHLTR	O84301 chlanydia t
10	46	32.9	658	1 KPCL_LYTPPI	Q25378 lytechinus
11	46	32.9	705	1 PURL_PYRHO	Q59621 pyrococcus
12	45	32.1	172	1 EPC_SYNY3	P73935 mycobactery
13	45	32.1	294	1 ECHC_MYCLE	P53526 mycobacteri
14	45	32.1	907	1 NUOG_ECOLI	P33602 escherichia
15	45	32.1	907	1 NUOG_SALTY	P33900 salmonella
16	45	32.1	925	1 NPEC2_HUMAN	Q13469 homo sapien
17	44.5	31.8	62	1 STAT_HUMAN	P02808 homo sapien
18	44.5	31.8	222	1 AMYB_SECC	P30271 secale cere
19	44.5	31.8	335	1 Y094_CAEEL	P41844 caenorhabdi
20	44.5	31.8	1056	1 YNN2_YEAST	P53914 saccharomyc
21	44	31.4	55	1 COX4_CERAP	O46589 cebus apell
22	44	31.4	137	1 COX4_PERPO	O46586 perodicticu
23	44	31.4	144	1 COX4_AOTAZ	O46584 aotus azara
24	44	31.4	169	1 COX4_BOVIN	P00423 bos taurus
25	44	31.4	169	1 COX4_RABIT	Q9ttt8 oryctolagus
26	44	31.4	282	1 IF34_SCHPO	P78795 schizosacch
27	44	31.4	587	1 ASO_CUCSA	P14133 cucumis sat
28	44	31.4	886	1 DSCI_MOUSE	P55849 mus musculu
29	44	31.4	1058	1 PMAL_DICDI	P54679 dictyosteli
30	43	30.7	54	1 IOVO_CIRAE	P05579 circos aeru
31	43	30.7	54	1 IOVO_GYPCO	P05578 gypts coprot
32	43	30.7	54	1 IOVO_HALAL	P52268 haliaetus
33	43	30.7	54	1 IOVO_HALIN	P05577 haliastur i

ALIGNMENTS

RESULT 1

ID	EPC_MOUSE	STANDARD;	PRT;	421 AA.
AC	P06336; P01855;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Ig epsilon chain C region.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84236092; PubMed=6329728;			
RA	Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;			
RA	"The nucleotide sequence of the mouse immunoglobulin epsilon gene:			
RT	comparison with the human epsilon gene sequence.";			
RL	EMBO J. 1:1117-1123(1982).			
RN	[2]			
RP	REVISIONS.			
RA	Honjo T.;			
RL	Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RX	SEQUENCE OF 34-421 FROM N.A.			
RA	Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;			
RT	"Cloning and nucleotide sequence of mouse immunoglobulin epsilon			
RT	chain cDNA.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X01857; CAA25977.1; -			
DR	EMBL; X01857; CAA25978.1; -			
DR	PIR; A02145; EHMSS.			
DR	PIR; A02144; EHMS.			
DR	HSSP; P01854; IIGF.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR003597; Ig_C1.			
DR	InterPro; IPR003600; Ig_Like.			
DR	Pfam; PF00047; Ig; 4.			
DR	SMART; SM00410; IG_Like; 2.			
DR	SMART; SM00407; IGC1; 2.			
DR	PROSITE; PS00290; IG_MHC; 3.			
KW	Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.			
FT	NON_TER 1 90			CH1.
FT	DOMAIN 1 90			CH2.
FT	DOMAIN 91 197			CH3.
FT	DOMAIN 198 304			

O46590 saimiri ust
O46582 saimiri sci
P08765 bacterioph
Q38039 bacterioph
P03637 bacterioph
P03638 bacterioph
O46333 pseudomonas
Q04633 pseudomonas
P53100 saccharomyc
P01854 homo sapien
Q00854 gallus gall
P48549 homo sapien
P35562 mus musculu

34 43 30.7 55 1 COX4_SAISC
35 43 30.7 124 1 COX4_SAISC
36 43 30.7 150 1 VGD_BPALK3
37 43 30.7 150 1 VGD_BPPHX
38 43 30.7 151 1 VGD_BPPHX
39 43 30.7 152 1 VGD_BPG4
40 43 30.7 182 1 APT_PSEAF
41 43 30.7 379 1 YGT5_YEAST
42 43 30.7 428 1 EPC_HUMAN
43 43 30.7 492 1 IRK3_CHICK
44 43 30.7 501 1 IRK3_HUMAN
45 43 30.7 501 1 IRK3_MOUSE


```
DB 513 YNEVDPRVKPTPVAGI 530
RESULT 4
MK32_YEAST
ID MK32_YEAST STANDARD; PRT; 363 AA.
AC P23060;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MAK32 protein.
GN MAK32 OR YCR019W OR YCR19W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE=89131254; PubMed=3916862;
RA Toh-E A., Sahashi Y.;
RT "The PEP18 locus of Saccharomyces cerevisiae: a complex locus
containing multiple genes.";
RL Yeast 1:159-171(1985).
RN [2]
RX SEQUENCE FROM N.A.
RP Feldmann H., Mannhaupt G., Vetter I.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NECESSARY FOR THE STRUCTURAL STABILITY OF L-A DOUBLE-
STRANDED RNA-CONTAINING PARTICLES. NECESSARY FOR GROWTH AT 37
DEGREES CELSIUS AS WELL AS FOR MAINTENANCE OF THE KILLER PLASMID.
CC -!- SIMILARITY: TO S.POMBE SPAC4G8.14C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DB DR EMBL; X59720; CAA42310.1; -
DR PIR; S19429; BVBYK2.
DR SGD; S0000612; MAK32.
DR InterPro: IPR002173; Pfkb.
DR Pfam: PF00294; pfkb; 1.
FT CONFLICT 15 15 I -> II (IN REF. 1).
FT CONFLICT 82 82 MISSING (IN REF. 1).
SQ SEQUENCE 363 AA; 40783 MW; CFB358F8ACF6EA4C CRC64;

Query Match 34.6%; Score 48.5; DB 1; Length 363;
Best Local Similarity 55.0%; Pred. No. 6.2;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 4 GYQSVIVDR-PDFPKPIVRSI 22
I : ||| |||| :|
DB 61 GLKWIIVDRGSDFFPKVIREI 80

RESULT 5
COX4_PITPI
ID COX4_PITPI STANDARD; PRT; 144 AA.
AC Q46595;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cytochrome c oxidase polypeptide IV (EC 1.9.3.1) (Fragment).
GN COX4.
OS Pithecia pithecia (White-faced saki).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Primates; Platyrrhini; Cebidae; Pitheciinae;
OC Pithecia.
OX NCBI_TaxID=43777;

Query Match 34.3%; Score 48; DB 1; Length 144;
Best Local Similarity 50.0%; Pred. No. 2.8;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 YGYQSVIVDRPDFPKPIVRSI 22
I : || ||| | :|
DB 8 YTLQSVIVDRRDYPLPDVAHV 27

RESULT 6
COX4_MOUSE
ID COX4_MOUSE STANDARD; PRT; 169 AA.
AC P19783;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome c oxidase polypeptide IV, mitochondrial precursor
(EC 1.9.3.1).
GN COX4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX SEQUENCE FROM N.A.
RP STRAIN=BALB/C; TISSUE=Heart;
RX MEDLINE=91057158; PubMed=2173832;
RA Grossman L.I., Akamatsu M.;
RT "Nucleotide sequence of a mouse cDNA for subunit IV of cytochrome c
oxidase.";
RL Nucleic Acids Res. 18:6454-6454(1990).
RN [2]
RX SEQUENCE FROM N.A.
RP MEDLINE=91378465; PubMed=1654830;
RA Carter R.S., Avadhani N.G.;
RT "Cloning and characterization of the mouse cytochrome c oxidase
subunit IV gene.";
RL Arch. Biochem. Biophys. 288:97-106(1991).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
MITOCHONDRIAL ELECTRON TRANSPORT.
CC NCBI_TaxID=10090;
CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
```

```
CC      c + 2 H(2)O.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M37831; AAB02139.1; -.
DR      EMBL; M58034; AAB02139.1; JOINED.
DR      EMBL; M37829; AAB02140.1; -.
DR      EMBL; X54691; CAA38507.1; -.
DR      PIR; S12142; S12142.
DR      PIR; S12508; S12508.
DR      PIR; S16114; S16114.
DR      HSSP; P00423; IOCC.
DR      MGD; MGI:88473; COX4.
DR      InterPro; IPR004203; COX4.
DR      Pfam; PF02936; COX4; 1.
DR      Oxidoreductase; Mitochondrion; Transit peptide.
KW      TRANSIT 1 22 MITOCHONDRION.
FT      CHAIN 23 169 CYTOCHROME C OXIDASE POLYPEPTIDE IV.
FT      CONFLICT 34 34 A -> R (IN REF. 1).
FT      CONFLICT 71 71 S -> N (IN REF. 1).
FT      CONFLICT 140 141 MQ -> IE (IN REF. 1; AAB02139 AND IN REF.
FT      2).
SQ      SEQUENCE 169 AA; 19530 MW; D30B1DBBE14FDBEA CRC64;

Query Match 34.3%; Score 48; DB 1; Length 169;
Best Local Similarity 36.4%; Pred. No. 3.3;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 3 YGYSIVDRPDPFKPIVRSITL 24
| : : | | | | | : |
DB 33 YAFPTYADRDYPLPDVAHVMT 54

RESULT 7
PTA_THEME STANDARD; PRT; 294 AA.
ID COX4_THEME
AC G9X0L4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphate acetyltransferase (EC 2.3.1.8) (Phosphotransacetylase).
GN PTA OR TM1130.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Uterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RL genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
RN [2]
RP SEQUENCE OF 1-39, AND CHARACTERIZATION.
RX MEDLINE=99173910; PubMed=10074080;
RA Bock A.-K., Glasenacker J., Schmidt R., Schoenheit P.;
RT "Purification and characterization of two extremely thermostable
RT enzymes, phosphate acetyltransferase and acetate kinase, from the
RT hyperthermophilic eubacterium Thermotoga maritima.";
RL J. Bacteriol. 181:1861-1867(1999).
```

```
CC      -!- FUNCTION: IN ADDITION TO ACETYL-COA (100%), THE ENZYME ACCEPTS
CC      PROPIONYL-COA (60%) AND BUTYRYL-COA (30%). HAS A TEMPERATURE
CC      OPTIMUM AT 90 DEGREES CELSIUS.
CC      -!- CATALYTIC ACTIVITY: Acetyl-CoA + phosphate -> CoA + acetyl
CC      phosphate.
CC      -!- PATHWAY: LAST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
CC      ACETYL-COA.
CC      -!- SUBUNIT: HOMOTETRAMER.
CC      -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC      -!- SIMILARITY: BELONGS TO THE PHOSPHATE ACETYLTRANSFERASE AND
CC      BUTYRYLTRANSFERASE FAMILY. MORE SIMILAR TO PTB THAN PTA.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE001770; AAD36206.1; -.
DR      TIGR; TM1130; -.
DR      InterPro; IPR002505; PTA_PTB.
DR      Pfam; PF01515; PTA_PTB; 1.
KW      Transferase; Acyltransferase; Complete proteome.
FT      CONFLICT 11 11 R -> Y (IN REF. 2).
SQ      SEQUENCE 294 AA; 32093 MW; F4B98B3CAE120AFB CRC64;

Query Match 34.3%; Score 48; DB 1; Length 294;
Best Local Similarity 46.7%; Pred. No. 5.9;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21
| : : | | | | | | |
DB 122 SVMIEPDPFRPLIIS 136

RESULT 8
COX4_RAT STANDARD; PRT; 169 AA.
ID COX4_RAT
AC P10888;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide IV, mitochondrial precursor
DE (EC 1.9.3.1).
GN COX4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=89240039; PubMed=2541414;
RA Goto Y., Amuro N., Okazaki T.;
RT "Nucleotide sequence of cDNA for rat brain and liver cytochrome c
RT oxidase subunit IV.";
RL Nucleic Acids Res. 17:2851-2851(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=89296488; PubMed=2544859;
RA Gopalan G., Droste M., Kadenbach B.;
RT "Nucleotide sequence of cDNA encoding subunit IV of cytochrome c
RT oxidase from fetal rat liver.";
RL Nucleic Acids Res. 17:4376-4376(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90237079; PubMed=2159010;
RA Yamada M., Amuro N., Goto Y., Okazaki T.;
RT "Structural organization of the rat cytochrome c oxidase subunit IV
```


gene.";
J. Biol. Chem. 265:7687-7692(1990).
[4]
SEQUENCE FROM N.A.
MEDLINE-90326528; PubMed-2165254;
Amuro N., Yamada M., Goto Y., Okazaki T.;
"Complete nucleotide sequence of the gene encoding rat cytochrome c
oxidase subunit IV.";
Nucleic Acids Res. 18:3992-3992(1990).
[5]
SEQUENCE FROM N.A.
MEDLINE-91067442; PubMed-2174541;
Virbasius J.V., Scarpulla R.C.;
"The rat cytochrome c oxidase subunit IV gene family: tissue-specific
and hormonal differences in subunit IV and cytochrome c mRNA
expression.";
Nucleic Acids Res. 18:6581-6586(1990).
[6]
SEQUENCE OF 23-32
STRAIN-WISTAR; TISSUE=Liver;
MEDLINE-95324529; PubMed-7601105;
Schaeffer H., Noack H., Halanek W., Brandt U., von Jagow G.;
"Cytochrome-c oxidase in developing rat heart. Enzymic properties and
amino-terminal sequences suggest identity of the fetal heart and the
adult liver isoform.";
Eur. J. Biochem. 230:235-241(1995).
-1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
MITOCHONDRIAL ELECTRON TRANSPORT.
-1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
c + 2 H(2)O.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X14209; CAA32426.1; -
EMBL; X15029; CAA33133.1; -
DBL; J05425; AAA40949.1; -
EMBL; X54081; CAA38018.1; -
PIR; A35209; A35209.
PIR; S04070; S04070.
PIR; S12724; S12724.
PIR; S14190; S14190.
HSP; P00423; LOCC.
InterPro: IPR004203; COX4.
Pfam; PF02936; COX4; 1.
Oxidoreductase; Mitochondrion; Transit peptide.
TRANSIT 1 22 MITOCHONDRION.
CHAIN 23 169 CYTOCHROME C OXIDASE POLYPEPTIDE IV.
SEQUENCE 169 AA; 19514 MW; F0187C76B7A1A9FE CRC64;
SQ

```

Query Match      33.6%; Score 47; DB 1; Length 169;
Best Local Similarity 45.5%; Pred. No. 4.7;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY      3 YGYQSIVDPRDPFKPIVRSITL 24
      | | | | | : | | : |
Db      33 YALPSYVDRRDYPLPDVAHVKL 54

RESULT      9
HEM3_CHLTR
ID      HEM3_CHLTR      STANDARD;      PRT;      241 AA.
AC      084301;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)

```

DE		Probable porphobilinogen deaminase (EC 4.3.1.8) (PBG)
DE		(hydroxymethylbilane synthase) (HMBS) (Pre-uroporphyrinogen synthase).
GN	HMC OR CT299.	
OS	Chlamydia trachomatis.	
OC	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.	
OX	NCBI_TaxID=813;	
[1]		
RN	SEQUENCE FROM N.A.	
RP	STRAIN=D/UW-3/CX;	
RC	MEDLINE=99000809; PubMed=9784136;	
RX	Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,	
RA	Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,	
RA	Davis R.W.;	
"	Genome sequence of an obligate intracellular pathogen of humans:	
RT	Chlamydia trachomatis.";	
RL	Science 282:754-759(1998).	
CC	-I- FUNCTION: TETRAPOLYMERIZATION OF THE MONOPYRROLE PBG INTO THE	
CC	HYDROXYMETHYLBILANE PREUPORPHYRINOGEN IN SEVERAL DISCRETE STEPS	
CC	(BY SIMILARITY).	
CC	-I- CATALYTIC ACTIVITY: 4 porphobilinogen + H(2O) =	
CC	hydroxymethylbilane + 4 NH(3).	
CC	-I- PATHWAY: FOURTH STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.	
CC	-I- SIMILARITY: BELONGS TO THE HMBS FAMILY.	

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	

DR	EMBL: AE001302; AAC67892.1; .	
DR	HSSP: P06983; IAH5	
DR	InterPro: IPR000860; Porphobil_deam.	
DR	Pfam: PF01379; Porphobil_deam; 1.	
DR	PROSITE; PS00533; PORPHOBILINOGEN_DEAM; FALSE_NEG.	
KW	porphyrin biosynthesis; Lyase; Complete proteome.	
SQ	SEQUENCE 241 AA; 26996 MW; F89FE0C20BF7044 CRC64;	
	Query Match 33.6%; Score 47; DB 1; Length 241;	
	Best Local Similarity 47.8%; Pred. No. 6.8;	
	Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;	
OY	1 CGVGYQSIVDPDPFKPIVRISIT 23	
Db	90 CDLGHSKADLPENPKATVVVSIT 112	
RESULT 10		
KPC1_LVTPI		
ID	KC1_LYTPI STANDARD; PRT; 658 AA.	
AC	Q25378;	
DT	15-JUL-1999 (Rel. 38, Created)	
DD	15-JUL-1999 (Rel. 38, Last sequence update)	
DE	15-JUL-1999 (Rel. 38, Last annotation update)	
DE	Protein kinase C (EC 2.7.1.-).	
GN	PKC1.	
OS	Lyttechinus pictus (Painted sea urchin).	
OS	Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;	
OC	Echinoidea; Euarchinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;	
OC	Lyttechinus.	
OX	NCBI_TaxID=7653;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE-Ovary;	
RA	Rakow T.I., Shen S.S.;	
RT	"Molecular cloning and characterization of protein kinase C from the	
RT	sea urchin Lyttechinus pictus.";	
RL	Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.	
CC	-I- FUNCTION: THIS IS A CALCIUM-ACTIVATED, PHOSPHOLIPID-DEPENDENT,	
CC	SERINE- AND THREONINE-SPECIFIC ENZYME (BY SIMILARITY).	

RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: CATALYZES A SALVAGE REACTION RESULTING IN THE FORMATION
 CC OF AMP, THAT IS ENERGICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.
 CC -1- CATALYTIC ACTIVITY: AMP + diphosphate = adenine + 5-phospho-
 CC alpha-D-ribose 1-diphosphate.
 CC -1- PATHWAY: PURINE SALVAGE
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE
 CC PHOSPHORIBOSYLTRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D90911; BAA18001.1; -;
 DR InterPro: IPR000836; PriBosyltran.
 DR InterPro: IPR002375; Pur_Pyr_pr_transf.
 DR Pfam: PF00156; PriBosyltran; 1.
 DR PROSITE: PS00103; PUR_PYR_PR_TRANSFER; 1.
 DR Transferase: Glycosyltransferase; Purine salvage; Complete proteome.
 KW SEQUENCE 172 AA; 18997 MW; D577E94BCE4E457 CRC64;
 SQ -----
 Query Match 32.1%; Score 45; DB 1; Length 172;
 Best Local Similarity 61.1%; Pred. No. 9.6;
 Matches 11; Conservative 1; Mismatches 4; Indels 2; Gaps 1;
 QY 8 IVDRPDPKP--IVRSIT 23
 Db | | | | | | | | | |
 Db 7 IRIDPDPKPFGIMPRDIT 24
 RESULT 13
 ECHC_MYCLE STANDARD; PRT; 294 AA.
 AC P53526;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable enoyl-CoA hydratase echa12 (EC 4.2.1.17).
 GN ECHA12 OR ML1241 OR MLCB1610.01 OR B1170_C2_224.
 OS Mycobacterium leprae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RA Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Holroyd S.M., Horsley K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Hain R., Hensby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodhead J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RT Nature 409:1007-1011(2001).
 RL -1- FUNCTION: COULD POSSIBLY OXIDIZES FATTY ACIDS USING SPECIFIC
 CC COMPONENTS (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxyacyl-CoA -> trans-2(or 3)-enoyl-
 CC CoA + H2(O).
 CC -1- SIMILARITY: BELONGS TO THE ENOYL-CoA HYDRATASE/ISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U00010; AAA17070.1; -;
 DR EMBL: AL049913; CAB43147.1; -;
 DR EMBL: AL583921; CAC31622.1; -;
 DR HSP: P14604; 2DUB.
 DR Leprosin; ML1241; -;
 DR InterPro: IPR001753; Enoyl_CoA_hydrase.
 DR Pfam: PF00378; ECH; 1.
 DR PROSITE: PS00166; ENOYL-CoA-HYDRATASE; 1.
 KW Fatty acid metabolism; Lyase; Complete proteome.
 SQ SEQUENCE 294 AA; 31491 MW; 0B7502E715A7D692 CRC64;
 Query Match 32.1%; Score 45; DB 1; Length 294;
 Best Local Similarity 50.0%; Pred. No. 17;
 Matches 11; Conservative 3; Mismatches 6; Indels 2; Gaps 1;
 QY 5 YQIVDRP--DFPKPIVRSITL 24
 Db | | | | | | | | | |
 Db 15 YRSVTLWLDFPRPEVALITL 36
 RESULT 14
 NUOG_ECOLI STANDARD; PRT; 907 AA.
 AC P33602; P76489; P78184; P78185;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE NADH dehydrogenase I chain G (EC 1.6.5.3) (NADH-ubiquinone
 DE oxidoreductase chain G) (NUO7).
 GN NUO7 OR B2283.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RA "The complete genome sequence of Escherichia coli K-12.";
 RT Science 277:1453-1474(1997).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97349980; PubMed=9205837;
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
 RA Oshima T., Oyama S., Saito N., Saito Y., Sato Y., Sivasubram S.,
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 RA Yamagata S., Horiuchi T.;
 RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
 RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
 RL analysis of its sequence features.";
 RL DNA Res. 4:91-113(1997).
 RN [3]
 RN SEQUENCE FROM N.A.

RC STRAIN=K12 / AN387;
RX MEDLINE=93389724; PubMed=7690854;
RA Weidner U., Geier S., Ptock A., Friedrich T., Leif H., Weiss H.;
RT "The gene locus of the proton-translocating NADH: ubiquinone
oxidoreductase in Escherichia coli. Organization and subunits of
relationship between the derived proteins and subunits of
mitochondrial complex I.";
RL J. Mol. Biol. 233:109-122(1993).
RN [4]
RN SEQUENCE OF 1-169 FROM N.A.
RP MEDLINE=94209210; PubMed=8157582;
RX Pruss B.M., Nelms J.M., Park C., Wolfe A.J.;
RA "Mutations in NADH:ubiquinone oxidoreductase of Escherichia coli
affect growth on mixed amino acids.";
RL J. Bacteriol. 176:2143-2150(1994).
RN [5]
RN SEQUENCE OF 1-12.
RP STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
TO BE UBIQUINONE. DOES COUPLE THE REDOX REACTION TO PROTON
TRANSLLOCATION AND THUS CONSERVES THE REDOX ENERGY IN A PROTON
GRADIENT.
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- COFACTOR: BINDS 1 2FE-2S CLUSTER AND 1 4FE-4S CLUSTER (Potential).
CC -1- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNITS NUOCD, E,
F, AND, G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 75 kDa SUBUNIT FAMILY.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
FRAMESHIFTS IN POSITIONS 714 AND 805.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000317; AAC75343.1; ALT_INIT.
DR EMBL; D90859; BAA16111.1; ALT_INIT.
DR EMBL; D90860; BAA16116.1; ALT_INIT.
DR EMBL; X68301; CAA48366.1; ALT_FRAME.
DR EMBL; L25055; AAA03538.1; -.
DR PIR; S37064; S37064.
DR PIR; S38316; S38316.
DR EcoGene; EG12087; nuog.
DR InterPro; IPR000283; Complex1_75k.
DR InterPro; IPR001041; Ferredoxin.
DR Pfam; PF00111; fer2; 1.
DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S;
KW Complete proteome.
FT INIT_MET 0
FT METAL 22 22 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 33 33 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 44 44 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 47 47 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 105 105 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 111 111 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 150 150 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 153 153 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 156 156 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 200 200 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 187 187 T -> Q (IN REF. 3).
FT CONFLICT 187

FT CONFLICT 209 209 T -> K (IN REF. 3).
FT CONFLICT 389 389 MISSING (IN REF. 3).
FT CONFLICT 647 647 S -> T (IN REF. 3).
SQ SEQUENCE 907 AA; 100168 MW; 4C258FEEF36533F2 CRC64;

Query Match 32.1%; Score 45; DB 1; Length 907;
Best Local Similarity 57.1%; Pred. No. 56;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYGQSVIVDRPDPF 15
 ||||: ||| |
Db 268 GYGYNLKDPRPQ 281

RESULT 15
NUOG_SALTY STANDARD; PRT; 907 AA.
AC P33900;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE NADH dehydrogenase I chain G (EC 1.6.5.3) (NADH-ubiquinone
oxidoreductase chain G) (NUO7).
GN NUOG OR STM2323 OR STY2553.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Ryan E., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 1-611 FROM N.A.
RC SPECIES=S.typhimurium;
RX MEDLINE=94052195; PubMed=8234329;
RA Archer C.D., Wang X., Elliott T.;
RT "Mutants defective in the energy-conserving NADH dehydrogenase of
Salmonella typhimurium identified by a decrease in energy-dependent
proteolysis after carbon starvation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9877-9881(1993).
RN [3]
RP SEQUENCE OF 611-907 FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=F98;
RX MEDLINE=98037521; PubMed=9371470;
RA Zhang-Barber L.Z., Turner A.K., Martin G., Fraenkel G., Dougan G.,
RA Barrow P.A.;
RT "Influence of genes encoding proton-translocating enzymes on
suppression of Salmonella typhimurium growth and colonization.";
RL J. Bacteriol. 179:7186-7190(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella

enterica serovar Typhi Ctl8.";
RT Nature 413:848-852(2001).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE. DOES COUPLE THE REDOX REACTION TO PROTON
CC TRANSLOCATION AND THUS CONSERVES THE REDOX ENERGY IN A PROTON
CC GRADIENT.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- COFACTOR: BINS 1 2FE-2S CLUSTER AND 1 4FE-4S CLUSTER (Potential).
CC -!- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNITS NUOCD, E,
CC F, AND, G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 75 kDa SUBUNIT FAMILY.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
CC FRAMESHIFTS IN POSITIONS 714 AND 805.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008804; AAL21224.1; ALT_INIT.
DR EMBL; L22504; AAL16063.1; ALT_INIT.
DR EMBL; L42521; -: NOT_ANNOTATED.CDS.
DR EMBL; AL627274; CAD07555.1; ALT_INIT.
DR ScyGene; SGI0257; nuog.
DR InterPro; IPR000283; Complex1_75K.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR001467; Molybdopterin.
DR Pfam; PF00111; fer2; 1.
DR Pfam; PF00384; molybdopterin; 1.
DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S;
KW Complete proteome.
FT INIT_MET 0 0 BY SIMILARITY.
FT METAL 22 22 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 33 33 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 44 44 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 47 47 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 102 102 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 105 105 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 111 111 IRON-SULFUR (2FE-2S) (POTENTIAL).
FT METAL 150 150 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 153 153 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 156 156 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 200 200 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT CONFLICT 677 677 A -> V (IN REF. 3).
FT CONFLICT 785 785 I -> L (IN REF. 3).
FT CONFLICT 846 846 A -> T (IN REF. 3).
FT CONFLICT 853 853 A -> T (IN REF. 3).
FT CONFLICT 887 887 M -> I (IN REF. 3).
SQ SEQUENCE 907 AA; 99896 MW; 5DE802178260C65E CRC64;

Query Match 32.1%; Score 45; DB 1; Length 907;
Best Local Similarity 57.1%; Pred. No. 56;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 GYGYQSIIVDRPDPF 15
 |||||::||| |
Db 268 GYGVNLKDRPROP 281

Search completed: August 26, 2002, 10:17:03
Job time: 699 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:16:19 ; Search time 66.82 seconds
(without alignments)
64.724 Million cell updates/sec

Title: US-09-701-623c-8
Perfect score: 140
Sequence: 1 CGYQSVDRDPKPKIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phage:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.5	37.5	460	10 Q94D33	Q94D33 oryza sativ
2	51.5	36.8	186	5 Q27123	Q27123 urechis cau
3	49	35.0	165	1 Q93701	Q93701 sulfolobus
4	49	35.0	433	16 Q9RTJ1	Q9RTJ1 deinococcus
5	49	35.0	149	5 Q9V917	Q9V917 drosophila
6	48.5	34.6	722	13 Q9PS25	Q9PS25 lampetra ja
7	48	34.3	753	16 Q9BMD1	Q9BMD1 rhizobium l
8	48	34.3	977	13 Q91925	Q91925 xenopus lae
9	48	34.3	1259	10 Q40983	Q40983 pisum sativ
10	47	33.6	534	12 Q65856	Q65856 beet yellow
11	47	33.6	591	16 Q92UW6	Q92UW6 rhizobium m
12	46.5	33.2	182	5 Q9V1Q8	Q9V1Q8 drosophila
13	46.5	33.2	210	16 Q9RUH0	Q9RUH0 deinococcus
14	46.5	33.2	711	10 Q9FGG5	Q9FGG5 arabidopsis
15	46	32.9	206	2 Q93AF4	Q93AF4 escherichia
16	46	32.9	351	16 Q69727	Q69727 mycobacteri

17	46	32.9	376	5 Q94259	Q94259 caenorhabdi
18	46	32.9	715	3 Q12204	Q12204 saccharomyc
19	46	32.9	870	16 Q98101	Q98101 rhizobium l
20	45.5	32.5	377	10 Q94GF6	Q94GF6 oryza sativ
21	45	32.1	273	3 Q96WV1	Q96WV1 schizosacch
22	45	32.1	536	10 Q94BP2	Q94BP2 arabidopsis
23	45	32.1	1354	13 Q9W6B2	Q9W6B2 xenopus lae
24	44.5	31.8	308	2 Q9AN61	Q9AN61 bradyrhizob
25	44.5	31.8	308	16 Q98NW9	Q98NW9 rhizobium l
26	44.5	31.8	330	16 Q9CB84	Q9CB84 mycobacteri
27	44.5	31.8	366	10 Q9ATS0	Q9ATS0 bothriochlo
28	44.5	31.8	392	5 Q9V883	Q9V883 drosophila
29	44.5	31.8	467	9 Q9AZ43	Q9AZ43 bacterioph
30	44.5	31.8	757	5 Q9VNF8	Q9VNF8 drosophila
31	44.5	31.8	769	5 Q9NEN9	Q9NEN9 drosophila
32	44	31.4	46	15 Q88403	Q88403 simian t-ce
33	44	31.4	116	2 Q9L4B7	Q9L4B7 uncultured
34	44	31.4	116	2 Q9L4A5	Q9L4A5 uncultured
35	44	31.4	182	11 Q9RLT0	Q9RLT0 mus musculu
36	44	31.4	222	4 Q9BUX1	Q9BUX1 homo sapien
37	44	31.4	699	17 Q9YFR1	Q9YFR1 aeropyrum p
38	44	31.4	706	16 Q9HWD2	Q9HWD2 pseudomonas
39	44	31.4	767	5 Q15882	Q15882 trypanosoma
40	44	31.4	963	2 Q9XCF5	Q9XCF5 mycobacteri
41	44	31.4	1213	5 Q9VE06	Q9VE06 drosophila
42	44	31.4	1219	5 Q9Y1A8	Q9Y1A8 drosophila
43	43.5	31.1	143	16 Q9X0G9	Q9X0G9 thermotoga
44	43.5	31.1	257	16 P72628	P72628 synechocyst
45	43.5	31.1	305	5 Q10667	Q10667 caenorhabdi

ALIGNMENTS

RESULT 1

Q94D33 ID Q94D33 PRELIMINARY; PRT; 460 AA.
AC Q94D33;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE P0712E02.23 PROTEIN.
GN P0712E02.23
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone: P0712E02.23";
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AP003492; BAB61859.1; -;
SQ SEQUENCE 460 AA; 49420 MW; 549DD3183A0408B6 CRC64;

Query Match 37.5%; Score 52.5; DB 10; Length 460;
Best Local Similarity 37.9%; Pred. No. 8.2;
Matches 11; Conservative 6; Mismatches 7; Indels 5; Gaps 1;

Qy 2 GYG-----YQSVDRDPKPKIVRSITLC 25
| | | | | : | | : | |
Db 266 GHGVFPVYSSMNSKDFPKVLLISLVLC 294

RESULT 2

Q27123 ID Q27123 PRELIMINARY; PRT; 186 AA.
AC Q27123;
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT IV.
OS Urechis caupo (Innkeeper worm) (Spoonworm).
OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.
OX NCBI_TaxID=6431;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94155469; PubMed=8111976;
RA Rosenthal E.;
RT "Sequence analysis of translationally controlled maternal mRNAs from
RL Urechis caupo.";
RL Dev. Genet. 14:485-491(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Rosenthal E.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U30468; AAA4396.1; -.
DR InterPro; IPR004203; COX4.
DR Pfam; PF02936; COX4; 1.
SQ SEQUENCE 186 AA; 21084 MW; CC8EC1FA5F84D766 CRC64;

Query Match 36.8%; Score 51.5; DB 5; Length 186;
Best Local Similarity 45.3%; Pred. No. 4.5;
Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 2 GYGYSI---VDRDPFPKPIVR 20
::: |::|::|::|
Db 48 GYGMNGLPYIDLPFPAPAVR 69

RESULT 3
O93701 PRELIMINARY; PRT; 165 AA.
AC O93701;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 19.2 KDA PROTEIN.
OS Sulfolobus sp. NOB8H2.
OC Plasmid pNOB8.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=84600;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB8H2;
RX MEDLINE=99044580; PubMed=9827331;
RA She Q., Phan H., Garrett R.A., Albers S.V., Stedman K.M., Zillig W.;
RT "Genetic profile of pNOB8 from Sulfolobus: the first conjugative
RL plasmid from an archaeon.";
RL Extremophiles 2:417-425(1998).
DR EMBL; AJ010405; CRA09149.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 165 AA; 19189 MW; 22E6A92532874A80 CRC64;

Query Match 35.0%; Score 49; DB 1; Length 165;
Best Local Similarity 58.8%; Pred. No. 9.5;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 YQSTVDRPFPKPIVRS 21
::: |::|::|::|
Db 71 YERIDKPDNPRIVRS 87

RESULT 4
O9RTJ1 PRELIMINARY; PRT; 433 AA.
AC O9RTJ1
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR1773.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1";
RL Science 286:1571-1577(1999).
DR EMBL; AE002019; AAF11328.1; -.
DR TIGR; DR1773; -.
DR InterPro; IPR001993; Mitoch.carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 433 AA; 46392 MW; 9089C90957FE4465 CRC64;

Query Match 35.0%; Score 49; DB 16; Length 433;
Best Local Similarity 35.3%; Pred. No. 27;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 6 QSVDRPDPFPKPIVRSI 22
::: |::|::|::|
Db 117 EAVAQRDLPEPLRLQ 133

RESULT 5
O9V917 PRELIMINARY; PRT; 1449 AA.
AC O9V917;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE CG16868 PROTEIN.
GN CG16868.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003792; AAF57485.1; -.
 DR FlyBase; FBgn0034498; CGI16868.
 DR InterPro; IPR004010; Cache.
 DR InterPro; IPR002035; vWFA.
 DR Pfam; PF02743; Cache; 3.
 DR PROSITE; PS50234; vWFA: 1.
 SQ SEQUENCE 1449 AA; 162090 MW; 70E3A155C4185D74 CRC64;

Query Match 35.0%; Score 49; DB 5; Length 1449;
 Best Local Similarity 33.3%; Pred. No. 97;
 Matches 10; Conservative 6; Mismatches 6; Indels 8; Gaps 1;

QY 2 GYGQSVIVDR-----PDPKPIVRSIT 23
 | | | : | | | | | : | : | : |
 Db 609 GLGYAFLLDRSTNGTLAHPAPRPRIQRET 638

RESULT 6
 Q9PSZ5 PRELIMINARY; PRT; 722 AA.
 AC Q9PSZ5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MANNOSYL-BINDING LECTIN-ASSOCIATED SERINE PROTEASE.
 OS Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 OC Petromyzontiformes; Petromyzontidae; Lethenteron.
 OX NCBI_TaxID=94989;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=99008558; PubMed=9794427;
 RA Endo Y., Takahashi M., Nakao M., Saiga H., Sekine H., Matsushita M.,
 RA Nonaka M., Fujita T.;
 RT "Two lineages of mannose-binding lectin-associated serine protease
 (MASP) in vertebrates.";
 RL J. Immunol. 161:4924-4930(1998).
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC EMBL; AB009075; BA86868.1; -.
 DR HSP; P00763; IDPO.
 DR MEROPS; S01.00A; -.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR InterPro; IPR001254; Trypsin.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00084; sushi; 2.

DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00032; CCP; 2.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00179; EGF_CA: 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01187; EGF_CA; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM: 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 DR Calcium-binding; EGF-like domain; Glycoprotein; Hydrolase; Lectin;
 KW Protease; Repeat; Serine protease.
 SQ SEQUENCE 722 AA; 79591 MW; 81DFB4F78F540B20 CRC64;

Query Match 34.6%; Score 48.5; DB 13; Length 722;
 Best Local Similarity 25.8%; Pred. No. 55;
 Matches 11; Conservative 4; Mismatches 7; Indels 21; Gaps 1;

QY 1 CGYGY-----QSIVDRDPFKPIVRSI 22
 | | | | | : | | | | | : | : | : |
 Db 177 CGYGYVLADNRCTVCSNHVFTERGELSSDPFPRYPAL 219

RESULT 7
 Q98MD1 PRELIMINARY; PRT; 753 AA.
 ID Q98MD1;
 AC Q98MD1;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DE DNA UPTAKE PROTEIN.
 DE DN MLL0627.
 GN MLL0627.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AF002995; BAB48182.1; -.
 KW Complete proteome.
 SQ SEQUENCE 753 AA; 80233 MW; C416697517B91EA9 CRC64;

Query Match 34.3%; Score 48; DB 16; Length 753;
 Best Local Similarity 53.3%; Pred. No. 68;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 GYQSVIVDRDPFKPI 18
 | | | | | : | | | | | : | : | : |
 Db 44 GYSLAVEPDPFAKPV 58

RESULT 8
 Q91925 PRELIMINARY; PRT; 977 AA.
 ID Q91925;
 AC Q91925;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE XTLD PROTEIN.

OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYO;
RA Lin J., Maeda R., Ong R., Kim J., Kung H., Maeno M.;
RT "Xtld, a Xenopus homolog of dorso-ventral polarity gene in Drosophila,
RT modifies tissue phenotypes of the ventral mesoderm.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: CONTAINS 5 CUB DOMAINS.
DR EMBL; D83476; BAAL1922.1; -.
DR HSSP; P00736; LAPO.
DR MEROPS; M12.005; -.
DR InterPro; IPR001506; Astacin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000859; CUB.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001130; Zn.MTpeptdse.
DR Pfam; PF01400; Astacin; 1.
DR Pfam; PF00431; CUB; 5.
DR Pfam; PF00008; EGF; 2.
DR PRINTS; PR00480; ASTACIN.
DR SMART; SM00042; CUB; 5.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00235; ZNMC; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01180; CUB; 5.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 977 AA; 110199 MW; 4D7D2E37C64FDF1F CRC64;

Query Match 34.3%; Score 48; DB 13; Length 977;
Best Local Similarity 47.1%; Pred. No. 90;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGYGYGIVDRPDPPKP 17
||| | : | | |
Db 908 CGDYNELYDGPDEAAP 924

RESULT 9
Q40983
ID Q40983 PRELIMINARY; PRT; 1259 AA.
AC Q40983;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE METALLOENDOPEPTIDASE.
OS Plasm sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95365331; PubMed=7638164;
RA VanderVerre P.S., Bennett T.M., Oblong J.E., Lamppa G.K.;
RT "A chloroplast processing enzyme involved in precursor maturation
RT shares a zinc-binding motif with a recently recognized family of
RT metalloendopeptidases.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:7177-7181(1995).
DR EMBL; U25111; AAA81472.1; -.
DR MEROPS; M16.004; -.
DR InterPro; IPR001431; Peptidase_M16.
DR Pfam; PF00675; Peptidase_M16; 1.
SQ SEQUENCE 1259 AA; 139573 MW; 5DBB33ED0D12383F CRC64;

Query Match 34.3%; Score 48; DB 10; Length 1259;
Best Local Similarity 38.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 5 YQSVDRPDPPKPPIVRSITLC 25
| | : | | : | | |
Db 79 YSSVLQOPTAPVPVRSQCTSC 99

RESULT 10
Q65856
ID Q65856 PRELIMINARY; PRT; 534 AA.
AC Q65856;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE P61.
OS Beet yellow stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Closterovirus
OX NCBI_TaxID=35290;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96266429; PubMed=8661428;
RA Karasev A.V., Nikolaeva O.V., Mushagian A.R., Lee R.F., Dawson W.O.;
RT "Organization of the 3'-terminal half of beet yellow stunt virus
RT genome and implications for the evolution of closteroviruses.";
RL Virology 221:199-207(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Karasev A.V., Nikolaeva O.V., Lee R.F., Wisler G.C., Duffus J.E.,
RA Dawson W.O.;
RT "Characterization of the beet yellow stunt virus coat protein gene.";
RL Phycopathology 88:1040-1045(1998).
DR EMBL; U51931; AAC55663.1; -.
SQ SEQUENCE 534 AA; 60979 MW; 65F0D86E486FE6F2 CRC64;

Query Match 33.6%; Score 47; DB 12; Length 534;
Best Local Similarity 42.4%; Pred. No. 68;
Matches 14; Conservative 1; Mismatches 6; Indels 12; Gaps 2;

QY 1 CGYGTARTARQVRVVRPDYELDGIFSKPIVMS 21
| | | | : | | | | | | | |
Db 357 CYGTGTARTARQVRVVRPDYELDGIFSKPIVMS 389

RESULT 11
Q92UW6
ID Q92UW6 PRELIMINARY; PRT; 591 AA.
AC Q92UW6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE ACETOLACTATE SYNTHASE ISOZYME I LARGE SUBUNIT PROTEIN
DE (EC 4.1.3.18).
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21966508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puchler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).

DR EMBL: AL603645; CAC49370.1; -.
KW Lyase; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 591 AA, 64157 MW, 7234D44FB221E597 CRC64;

Query Match 33.6%; Score 47; DB 16; Length 591;
Best Local Similarity 52.6%; Pred. No. 75;
Matches 10; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 1 CGYGYSIVDR----PDFF 15
||||| ||| |||
Db 433 CGYGFPSIVGAKIGCDVP 451

RESULT 12
Q9VIQ8 ID Q9VIQ8 PRELIMINARY; PRT; 182 AA.
AC Q9VIQ8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CG10664 PROTEIN (GMI4452P).
GN CG10664.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.E., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hosten D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,

RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003664; AAF53857.1; -.
DR EMBL: AY058491; AAL13720.1; -.
DR FlyBase: FBgn0032833; CG10664.
DR InterPro: IPR004203; COX4.
DR Pfam: PF02936; COX4; 1.
SQ SEQUENCE 182 AA; 20519 MW; 6D5953D8DE41217A CRC64;

Query Match 33.2%; Score 46.5; DB 5; Length 182;
Best Local Similarity 45.5%; Pred. No. 26;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 2 GYGYSQ---IVDRPDPKPIVR 20
||||| ||| ||| |||
Db 43 GYGWNGTACVADRVDPPLPAVR 64

RESULT 13
Q9RUH0 ID Q9RUH0 PRELIMINARY; PRT; 210 AA.
AC Q9RUH0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 23.0 KDA PROTEIN.
GN DR1416.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577 (1999).
DR EMBL: AE001986; AAF10992.1; -.
DR TIGR: DR1416; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 210 AA; 22978 MW; 326ECF600C867E6 CRC64;

Query Match 33.2%; Score 46.5; DB 16; Length 210;
Best Local Similarity 45.5%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 3 YGYQSIQVDRPDPKPIVRITL 24
:|||| ||| ||| |||
Db 171 FGYYQ---PRDRPAPLPSGLTL 189

RESULT 14
Q9FGG5 ID Q9FGG5 PRELIMINARY; PRT; 711 AA.
AC Q9FGG5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILARITY TO UNKNOWN PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; PubMed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones.";
 RL DNA Res. 7:31-63(2000).
 DR EMBL; AB025635; BAB08932.1; --
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR SMART; SM00220; S_TKG; 1.
 DR SMART; SM00219; TyKc; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; UNKNOWN_1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 711 AA; 81868 MW; AE31C91CA24F8AE0 CRC64;

Query Match 33.2%; Score 46.5; DB 10; Length 711;
 Best Local Similarity 33.3%; Pred. NO. 1.1e+02;
 Matches 8; Conservative 5; Mismatches 6; Indels 5; Gaps 1;

QY 1 CGYGYQSIV-----DRPDPFKPIV 19
 ||:| |:| |:| |:| |:|
 Db 382 CGFGLSSVVKQAQSSSKPGTPDPVI 405

RESULT 15
 Q93AF4
 ID Q93AF4 PRELIMINARY; PRT; 206 AA.
 AC Q93AF4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PUTATIVE OXIDOREDUCTASE.
 OS Escherichia coli.
 OG plasmid p0113.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OC NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sloan J., Hartland E.;
 RT "Putative transfer region (partial) and type IV pilus operon encoded
 RT by the large virulence plasmid, p0113, of enterohemorrhagic
 RT Escherichia coli O113:H21.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF422146; AAL18824.1; --
 KW plasmid.
 SQ SEQUENCE 206 AA; 22355 MW; DED02904AB24C87A CRC64;

Query Match 32.9%; Score 46; DB 2; Length 206;
 Best Local Similarity 43.8%; Pred. NO. 35;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFK 16
 ||:| |:| |:| |:| |:|
 Db 33 CGHSLSSAMDTPLDPE 48

Search completed: August 26, 2002, 10:16:21
 Job time: 712 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:07:02 ; Search time 82.75 Seconds
(without alignments)
33.557 Million cell updates/sec

Title: US-09-701-623c-8
Perfect score: 140
Sequence: 1 CGYGQSVDRDPKPIVRSLTLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	25	AAV80001	Optimised IgE-CH3
2	129	92.1	45	AAV80018	IgE immunogenic pe
3	129	92.1	46	AAV80017	IgE immunogenic pe
4	116	82.9	561	AAV17415	Mouse immunoglobul
5	113	80.7	313	AAV79997	Mouse immunoglobul
6	106	75.7	25	AAV80000	Optimised IgE-CH3
7	85	60.7	313	AAV79996	Rat immunoglobulin
8	85	60.7	340	AAV803643	Rat IgE heavy chai
9	85	60.7	341	AAV806206	Immunogenic peptid
10	74	52.9	25	AAV80077	Optimised IgE-CH3
11	66	47.1	25	AAV79999	Optimised IgE-CH3

12	66	47.1	45	21	AAV80019	IgE immunogenic pe
13	66	47.1	45	21	AAV80083	IgE immunogenic pe
14	66	47.1	46	21	AAV80020	IgE immunogenic pe
15	66	47.1	57	21	AAV80081	IgE immunogenic pe
16	66	47.1	62	21	AAV80080	IgE immunogenic pe
17	66	47.1	63	21	AAV80084	IgE immunogenic pe
18	64	45.7	25	21	AAV79998	Optimised IgE-CH3
19	64	45.7	25	21	AAV91212	Modified human IgE
20	64	45.7	25	21	AAV88602	Peptide sequence o
21	64	45.7	42	21	AAV80014	IgE immunogenic pe
22	64	45.7	42	21	AAV91215	Modified MvF Th ep
23	64	45.7	42	21	AAV91216	Modified MvF Th ep
24	64	45.7	42	21	AAV91217	Modified MvF Th ep
25	64	45.7	42	21	AAV68604	Peptide sequence o
26	64	45.7	45	21	AAV80007	IgE-CH3 domain ant
27	64	45.7	45	21	AAV91218	Modified HBV surfa
28	64	45.7	45	21	AAV68605	Peptide sequence o
29	64	45.7	46	21	AAV80011	IgE immunogenic pe
30	64	45.7	46	21	AAV91213	Modified MvF Th ep
31	64	45.7	46	21	AAV91214	Modified MvF Th ep
32	64	45.7	46	21	AAV68603	Peptide sequence o
33	64	45.7	56	21	AAV80016	IgE immunogenic pe
34	64	45.7	59	21	AAV80010	IgE immunogenic pe
35	64	45.7	60	21	AAV80013	IgE immunogenic pe
36	64	45.7	60	21	AAV80015	IgE immunogenic pe
37	64	45.7	63	21	AAV80008	IgE-CH3 domain ant
38	64	45.7	63	21	AAV80012	IgE immunogenic pe
39	64	45.7	63	21	AAV91219	Inv epitope/modifi
40	64	45.7	63	21	AAV68606	Peptide sequence o
41	63	45.0	60	21	AAV80078	IgE immunogenic pe
42	58	41.4	345	21	AAV806207	Immunogenic peptid
43	54	38.6	15	21	AAV50899	Antibody 15A.2 mur
44	53	37.9	17	21	AAV50898	Antibody 15A.2 swi
45	50	35.7	344	22	AAU04826	Micromonospora eve

ALIGNMENTS

RESULT 1

AAV80001
ID AAV80001 standard; Peptide: 25 AA.

XX AAV80001;

XX 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Mus sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX Claim 1; Page 100; 155pp; English.
 PS The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 7.9e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGYGYQSVDRPDPKPIVRSITLC 25
 Db 1 CGYGYQSVDRPDPKPIVRSITLC 25

RESULT 2
 AAY80018
 ID AAY80018 standard; Peptide; 45 AA.

XX AAY80018;
 XX 15-MAY-2000 (first entry)
 DT IgE immunogenic peptide conjugate SEQ ID NO:25.

DE Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 XX immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX

OS Unidentified.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfeld AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 76; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 45 AA;

Query Match 92.1%; Score 129; DB 21; Length 45;
 Best Local Similarity 92.0%; Pred. No. 7.8e-13;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYGYQSVDRPDPKPIVRSITLC 25
 Db 21 CGYGYQSVDRPDPKPIVRSITLC 45

RESULT 3
 AAY80017
 ID AAY80017 standard; Peptide; 46 AA.

XX AAY80017;

XX 15-MAY-2000 (first entry)

DE IgE immunogenic peptide conjugate SEQ ID NO:24.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX

OS Unidentified.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfeld AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 76; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
SQ Sequence 46 AA;

Query Match 92.1%; Score 129; DB 21; Length 46;
Best Local Similarity 92.0%; Pred. No. 8e-13;
Matches 23; Conservative 0; Mismatches 0; Gaps 0;

QY 1 CGYGYQIVDRPDPKPIVRSITL 25
Db 22 cgygyqivdhpdpkpivrstl 46
|||||

RESULT 4
AAY17415
ID AAY17415 standard; Protein; 561 AA.
XX
AC AAY17415;
XX
DT 26-JUL-1999 (first entry)
XX
DE Mouse immunoglobulin E heavy chain.
XX
KW Mouse; immunoglobulin E; IgE; allergy; transgenic animal.
XX
OS Mus sp.
XX

FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal
FT Protein 20..561
FT /label= immunoglobulin_E_heavy_chain
XX
PN EP921189-A1.

XX
XX 09-JUN-1999.
XX
XX 13-NOV-1998; 98EP-0309340.
XX
XX 14-NOV-1997; 97JP-0313989.
XX

XX (SANY) SANKYO CO LTD.
XX (TOKM-) TOKYO METROPOLITAN INST MEDICAL SCI.
XX
XX Karasuyama H, Matsuoka K, Taya C, Yonemawa H;
XX WPI; 1999-315404/27.
XX N-PSDB; AAX61084.
XX Transgenic non-human animal allergy models
XX
XX Claim 27; Page 24-27; 42pp; English.
XX

XX The present invention describes a transgenic, non-human animal with its
XX genome altered to constitutively express a molecule having a constant
XX region which can bind an IgE receptor on mast cells in the animals, the
XX molecule having an immunoglobulin structure and being further capable of
XX specifically binding a predetermined antigen. The transgenic animal is
XX useful as a model for evaluating the activity and the ability of
XX substances i.e. with antiallergic activity, to affect any allergic
XX reaction caused in the animal by the administration of the antigen to
XX the animal, and applying the substance to be evaluated. The present
XX sequence is the mouse immunoglobulin E (IgE) heavy chain, given in
XX the present invention.

SQ Sequence 561 AA;

Query Match 82.9%; Score 116; DB 20; Length 561;
Best Local Similarity 91.3%; Pred. No. 1.4e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 GYGVQIVDRPDPKPIVRSITL 24
Db 419 gyygvqvdrdpdpkpivrstl 441
|||||

RESULT 5
AAY79997
ID AAY79997 standard; Protein; 313 AA.
XX
AC AAY79997;
XX
DT 15-MAY-2000 (first entry)
XX
DE Mouse immunoglobulin E epsilon heavy chain SEQ ID NO:4.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Mus sp.
XX
XX WO9967293-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13959.
XX
XX 20-JUN-1998; 98US-0100287.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY, Walfield AM;
XX
XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Example 1; Page 66-68; 155pp; English.
XX
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils,
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX
XX Sequence 313 AA;

Query Match 80.7%; Score 113; DB 21; Length 313;
Best Local Similarity 95.5%; Pred. No. 2.1e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGVQIVDRPDPKPIVRSIT 23
Db 419 gyygvqvdrdpdpkpivrstl 441
|||||

Db 191 ggygqclvdprdpfpkplvrsit 212

RESULT 6
 AAY80000
 ID AAY80000 standard; Peptide; 25 AA.
 XX
 AC AAY80000;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for rat IgE.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Rattus sp.
 OS Synthetic.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX
 DR WPI; 2000-160578/14.
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX
 PS Claim 1; Page 99; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 75.7%; Score 106; DB 21; Length 25;
 Best Local Similarity 80.0%; Pred. No. 1.6e-09;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYGYSIVDRDPFPKPIVRSITLC 25
 II IIII I I IIIIIIIII I
 Db 1 cgygqsrldhphfpkplvrsitkc 25

RESULT 7
 AAY79996
 ID AAY79996 standard; Protein; 313 AA.

XX AAY79996;
 AC 15-MAY-2000 (first entry)
 DT
 DE Rat immunoglobulin E epsilon heavy chain SEQ ID NO:3.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Rattus sp.
 OS WO9967293-A1.
 PN
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX
 DR WPI; 2000-160578/14.
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX
 PS Example 1; Page 66-68; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 313 AA;

Query Match 60.7%; Score 85; DB 21; Length 313;
 Best Local Similarity 77.3%; Pred. No. 4.9e-05;
 Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGYSIVDRDPFPKPIVRSIT 23
 I III I I IIIIIIIII
 Db 190 ggygqsrldhphfpkplvrsit 211

RESULT 8
 AAB03643
 ID AAB03643 standard; protein; 340 AA.
 XX
 AC AAB03643;
 XX
 DT 22-NOV-2000 (first entry)
 XX
 DE Rat IgE heavy chain constant regions 2, 3 and 4.
 XX

KW Rat; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX
 OS Rattus sp.

XX
 XX WO200025722-A2.

XX PN 11-MAY-2000.

XX PD 21-OCT-1999; 99WO-SE01896.

XX PF 02-NOV-1998; 98US-0106652.

XX PR 22-SEP-1999; 99US-0401636.

XX XX (RESI-) RESISTENTIA PHARM AB.

XX PA Hellman LT;

XX PI WPI; 2000-365342/31.

XX DR Immunogenic polypeptides useful for preventing the harmful effects of
 immunoglobulin E in mammals -

XX PS Disclosure; Fig 1; 50pp; English.
 XX CC The present sequence is an immunogenic peptide consisting of the
 heavy chain constant regions 2, 3 and 4 of the rat IgE. It was used to
 construct a number of immunogenic peptides which consisted of regions of
 IgE from different mammals, which appear to cause a stronger polyclonal
 anti-self IgE response than peptides consisting of the same regions from
 one mammal. Immunogenic peptides, particularly those consisting of
 different heavy chain constant regions, can be used for vaccination
 in humans, against bacterial and viral infections and allergies, such
 as asthma, fur, pollen and food allergies and eczema.

XX XX Sequence 340 AA;

Query Match 60.7%; Score 85; DB 21; Length 340;
 Best Local Similarity 77.3%; Pred. No. 5.4e-05;
 Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGQSVIVDRPDPKPIVRSIT 23

Db 195 ggyqcrvdhphfpkpvrsit 216

RESULT 9

ID AAB06206

XX AAB06206 standard; protein; 341 AA.

AC AAB06206;

XX DT 22-NOV-2000 (first entry)

XX DE Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.

XX KW Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
 asthma; eczema; immunogenic peptide.

XX OS Chimeric - Didelphis virginiana.

XX OS Chimeric - Rattus sp.

XX PN WO200025722-A2.

XX PD 11-MAY-2000.

XX PF 21-OCT-1999; 99WO-SE01896.

XX PR 02-NOV-1998; 98US-0106652.

XX PR 22-SEP-1999; 99US-0401636.

XX XX (RESI-) RESISTENTIA PHARM AB.

XX PI Hellman LT;

XX DR WPI; 2000-365342/31.

XX XX Immunogenic polypeptides useful for preventing the harmful effects of
 immunoglobulin E in mammals -

XX PS Disclosure; Fig 2; 50pp; English.

XX CC The present sequence is an immunogenic peptide consisting of the heavy
 chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 constant region 3 from the rat. It was shown to cause a stronger
 polyclonal anti-self IgE response than peptides consisting of the same
 regions from one mammal. Immunogenic peptides, particularly those
 consisting of different heavy chain constant regions, can be used for
 vaccination in humans, against bacterial and viral infections and
 allergies, such as asthma, fur, pollen and food allergies and eczema.

XX XX Sequence 341 AA;

Query Match 60.7%; Score 85; DB 21; Length 341;
 Best Local Similarity 77.3%; Pred. No. 5.4e-05;
 Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGQSVIVDRPDPKPIVRSIT 23

Db 206 ggyqcrvdhphfpkpvrsit 227

RESULT 10

AY80077

ID AAY80077 standard; Peptide; 25 AA.

XX AC AAY80077;

XX DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 antibody; allergy; allergic disease; immunisation; anti-allergic;
 anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Equus caballus.

XX OS Synthetic.

XX PN WO967293-A1.

XX XX 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX XX (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 for immunization against allergy -
 XX PS Claim 1; Page 146; 155pp; English.
 XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 and anti-asthmatic properties. (I) induces polyclonal antibodies
 specific for a target effector site on the epsilon-heavy chain of IgE,
 and so preventing triggering and activation of mast cells and basophils

Qy	1	CGYGYQSIVDRPDPFKPIVRSITL	25
Db	21	cgetysrvthpdkpdivrsi	45
RESULT	14		
AA	AY80020		
ID	AA	Y80020 standard; Peptide; 46	AA.
AC	AA	Y80020;	
XX	XX		
XX	XX		
DT	15-MAY-2000	(first entry)	
XX	XX		
DE	DE	IgE immunogenic peptide conjugate	SEQ ID NO:27.
XX	XX		
KW	KW	Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;	
KW	KW	immunogenic; immunostimulatory; carrier protein; helper T cell epitope;	
KW	KW	antibody; allergy; allergic disease; immunisation; anti-allergic;	
KW	KW	anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.	
XX	XX		
OS	OS	Unidentified.	
XX	XX		
PN	PN	WO9967293-A1.	
XX	XX		
PD	PD	29-DEC-1999.	
XX	XX		
PF	PF	21-JUN-1999; 99WO-US13959.	
XX	XX		
PR	PR	20-JUN-1998; 98US-0100287.	
XX	XX		
PA	PA	(UNBI-) UNITED BIOMEDICAL INC.	
XX	XX		
PI	PI	Wang CY, Walfield AM;	
XX	XX		
DR	DR	WPI; 2000-160578/14.	
XX	XX		
PT	PT	New antigenic peptide from the CH3 domain of immunoglobulin E, fusions	
PT	PT	for Immunization against allergy	
XX	XX		
PS	PS	Claim 14; Page 77; 155pp; English.	
XX	XX		
CC	CC	The present invention describes immunoglobulin E (IgE)-CH3 domain	
CC	CC	antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic	
CC	CC	and anti-asthmatic properties. (I) induces polyclonal antibodies	
CC	CC	specific for a target effector site on the epsilon-heavy chain of IgE,	
CC	CC	and so preventing triggering and activation of mast cells and basophils	
CC	CC	and downregulation of IgE synthesis. Conjugates, or fusion peptides,	
CC	CC	containing (I) are used for active immunisation against IgE-mediated	
CC	CC	allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy	
CC	CC	dermatitis. Nucleic acids that encode these compounds are useful for	
CC	CC	recombinant production of corresponding peptides or in DNA vaccines.	
CC	CC	Conjugates of (I) that include a promiscuous T helper cell epitope	
CC	CC	(functional in genetically diverse subjects), in addition to a B cell	
CC	CC	target epitope, have increased immunogenicity and may include cyclic	
CC	CC	constraints (disulfide bridge) to stabilise conformational features and	
CC	CC	maximize cross-reactivity to the natural target. They induce safe	
CC	CC	(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino	
CC	CC	acid sequences used in the exemplification of the present invention.	
XX	XX		
SQ	SQ	Sequence	46 AA;
Query Match		47.1%;	Score 66; DB 21; Length 46;
Best Local Similarity		56.08;	Pred. No. 0.0053;
Matches 14; Conservative		0; Mismatches 11; Indels	0; Gaps 0;
Qy	1	CGYGYQSIVDRPDPFKPIVRSITL	25
Db	22	cgetysrvthpdkpdivrsi	46
RESULT	15		
AA	Y80081		

ID AAY80081 standard; Peptide; 57 AA.
XX AC
XX AAY80081;
XX DT
XX 15-MAY-2000 (first entry)
XX DE
XX IgE immunogenic peptide conjugate SEQ ID NO:88.
XX KW
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS
XX Unidentified.
XX PN
XX WO9967293-A1.
XX PD
XX 29-DEC-1999.
XX PF
XX 21-JUN-1999; 99WO-US13959.
XX PR
XX 20-JUN-1998; 98US-0100287.
XX PA
XX (UNBI-) UNITED BIOMEDICAL INC.
XX PI
XX Wang CY, Walfield AM;
XX DR
XX WPI; 2000-160578/14.
XX PT
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Claim 14; Page 77; 155pp; English.
XX
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX SQ
XX Sequence 57 AA;

Query Match 47.1%; Score 66; DB 21; Length 57;
Best Local Similarity 56.0%; Pred. No. 0.0067;
Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGYQSIIVDRPDPFKPIVRSITILC 25
||| | | | | | | | |
Db 33 cgetysrvthplpkdivrsiakc 57

Search completed: August 26, 2002, 10:07:02
Job time: 228 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:05:18 ; Search time 34.18 Seconds
(without alignments)
17.865 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 GYGQSVDRDPFKPIVRSITL 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	116	82.9	561	3	US-09-192-545-2
2	64	45.7	25	3	US-09-100-414B-95
3	64	45.7	25	4	US-09-303-323-95
4	64	45.7	42	3	US-09-100-414B-98
5	64	45.7	42	3	US-09-100-414B-99
6	64	45.7	42	3	US-09-100-414B-100
7	64	45.7	42	4	US-09-303-323-98
8	64	45.7	42	4	US-09-303-323-99
9	64	45.7	42	4	US-09-303-323-100
10	64	45.7	45	3	US-09-100-414B-101
11	64	45.7	45	4	US-09-303-323-101
12	64	45.7	46	3	US-09-100-414B-96
13	64	45.7	46	3	US-09-100-414B-97
14	64	45.7	46	4	US-09-303-323-96
15	64	45.7	46	4	US-09-303-323-97
16	64	45.7	63	3	US-09-100-414B-102
17	64	45.7	63	4	US-09-303-323-102
18	48	34.3	268	3	US-09-187-049-9
19	48	34.3	1259	3	US-09-187-049-13
20	45	32.1	426	1	US-08-336-583-2
21	45	32.1	426	5	PCT-US95-13795-2
22	45	32.1	699	5	PCT-US94-07297-39
23	45	32.1	921	1	US-08-396-479B-2
24	45	32.1	921	1	US-08-818-823-2
25	44.5	31.8	350	4	US-09-821-736-2
26	44.5	31.8	396	4	US-09-461-474-2
27	43	30.7	106	2	US-08-232-539D-54

28	43	30.7	107	4	US-09-187-859-33	Sequence 33, Appl
29	43	30.7	113	2	US-08-232-539D-56	Sequence 56, Appl
30	43	30.7	501	1	US-08-103-445-4	Sequence 4, Appl
31	43	30.7	501	1	US-08-461-690B-4	Sequence 2, Appl
32	43	30.7	501	1	US-08-486-342-2	Sequence 2, Appl
33	43	30.7	501	1	US-08-473-092-2	Sequence 2, Appl
34	43	30.7	501	1	US-08-614-801A-2	Sequence 2, Appl
35	43	30.7	501	1	US-08-066-371-2	Sequence 2, Appl
36	43	30.7	501	4	US-09-042-494-2	Sequence 2, Appl
37	43	30.7	1290	1	US-08-470-350B-2	Sequence 2, Appl
38	43	30.7	1843	4	US-09-413-814-50	Sequence 50, Appl
39	42	30.0	482	2	US-08-876-874-2	Sequence 2, Appl
40	41.5	29.6	248	3	US-08-750-145A-18	Sequence 18, Appl
41	41.5	29.6	248	3	US-08-975-698A-22	Sequence 22, Appl
42	41.5	29.6	248	4	US-09-417-090-22	Sequence 22, Appl
43	41.5	29.6	248	4	US-09-727-578-22	Sequence 41, Appl
44	41.5	29.6	776	2	US-08-671-978A-41	Sequence 11, Appl
45	41	29.3	1464	2	US-08-231-193A-11	

ALIGNMENTS

RESULT 1
US-09-192-545-2
; Sequence 2, Application US/09192545
; Patent No. 6118044
; GENERAL INFORMATION:
; APPLICANT: Karasuyama, Hajime
; APPLICANT: Yonekawa, Hiromichi
; APPLICANT: Taya, Choji
; TITLE OF INVENTION: Transgenic Animal Allergy Models and Methods for Their Use
; FILE REFERENCE: 799P79570
; CURRENT APPLICATION NUMBER: US/09/192,545
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: JP HEI 9-313989
; EARLIER FILING DATE: 1997-11-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Designed heavy
US-09-192-545-2

Query Match 82.9%; Score 116; DB 3; Length 561;
Best Local Similarity 91.3%; Pred. No. 7.7e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGQSVDRDPFKPIVRSITL 24
|||||:|||||:|||||:|||||:
Db 419 GYGQCVDRDPFKPIVRSITL 441

RESULT 2
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95
```

```
Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 CGGYQSIQSDRPPDPKPIVRSITLC 25
|| ||| | | | : || | |
Db 1 CGETYQSRVTHPHLPRALMRSTTKC 25
```

```
RESULT 3
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 45.7%; Score 64; DB 4; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 CGGYQSIQSDRPPDPKPIVRSITLC 25
|| ||| | | | : || | |
Db 1 CGETYQSRVTHPHLPRALMRSTTKC 25

RESULT 4
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98
```

```
Query Match 45.7%; Score 64; DB 3; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 CGGYQSIQSDRPPDPKPIVRSITLC 25
|| ||| | | | : || | |
Db 18 CGETYQSRVTHPHLPRALMRSTTKC 42

RESULT 5
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
```


; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 45.7%; Score 64; DB 3; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSVIVDRDPFKPIVRSITLC 25
|| ||| | | | :|| | |
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 6
US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849

; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 45.7%; Score 64; DB 3; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSVIVDRDPFKPIVRSITLC 25
|| ||| | | | :|| | |
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 7
US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSVIVDRDPFKPIVRSITLC 25
|| ||| | | | :|| | |
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8
US-09-303-323-99

; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDKPIVRSITLC 25
||| ||| | | | : || | |
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 9
US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDKPIVRSITLC 25
||| ||| | | | : || | |
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 10
US-09-100-414B-101
; Sequence 101, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-101

Query Match 45.7%; Score 64; DB 3; Length 45;

Best Local Similarity 48.0%; Pred. No. 0.004;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSIIVDRPDPKPIVRSITLC 25
|| ||| | | : || | |
Db 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 11

US-09-303-323-101
; Sequence 101, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-101

Query Match 45.7%; Score 64; DB 4; Length 45;
Best Local Similarity 48.0%; Pred. No. 0.004;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSIIVDRPDPKPIVRSITLC 25
|| ||| | | : || | |
Db 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 12

US-09-100-414B-96
; Sequence 96, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue

; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-96

Query Match 45.7%; Score 64; DB 3; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0041;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSIIVDRPDPKPIVRSITLC 25
|| ||| | | : || | |
Db 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 13

US-09-100-414B-97
; Sequence 97, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 45.7%; Score 64; DB 3; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0041;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYSIVDRPDPKPIVRSITLC 25
Db 22 CGETYQSRVTHPHLPALMRSTTKC 46

RESULT 14

US-09-303-323-96
; Sequence 96, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-96

Query Match 45.7%; Score 64; DB 4; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0041;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYSIVDRPDPKPIVRSITLC 25
Db 22 CGETYQSRVTHPHLPALMRSTTKC 46

RESULT 15

US-09-303-323-97
; Sequence 97, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-09-303-323-97

Query Match 45.7%; Score 64; DB 4; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0041;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYSIVDRPDPKPIVRSITLC 25
Db 22 CGETYQSRVTHPHLPALMRSTTKC 46

Search completed: August 26, 2002, 10:05:19
Job time: 125 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:04:23 ; Search time 45.87 Seconds
(without alignments)
52.370 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	88	62.4	426	2	I36948	Ig epsilon-chain - chimpanzee (fragment)
2	85	60.3	428	1	EHHU	Ig epsilon chain C
3	82	58.2	429	1	EHRT	Ig epsilon chain C
4	58	41.1	107	2	I68730	IgE chain C3 regio
5	58	41.1	107	2	I68726	IgE chain C3 regio
6	58	41.1	388	1	EHMS	Ig epsilon chain C
7	58	41.1	548	2	S38864	Ig epsilon chain C
8	54	38.3	116	2	S37909	Ig epsilon chain C
9	54	38.3	394	2	T19571	hypothetical prote
10	54	38.3	1306	2	S42659	hypothetical prote
11	54	38.3	1625	2	T02921	acetyl-CoA carboxy
12	54	38.3	2311	2	T06161	acetyl-CoA carboxy
13	54	38.3	2325	2	T02235	acetyl-CoA carboxy
14	52	36.9	1756	2	T02599	hypothetical prote
15	51.5	36.5	605	1	FPMS	alpha-fetoprotein
16	51	36.2	106	2	A03859	hypothetical prote
17	51	36.2	195	2	A75171	hypothetical prote
18	50	35.5	539	2	T48197	hypothetical prote
19	49.5	35.1	381	2	S28115	gas-vesicle protei
20	49.5	35.1	382	2	J01122	gas-vesicle protei
21	49.5	35.1	382	2	T08243	gas-vesicle operon
22	49	34.8	684	2	S60286	novel antigen rece
23	49	34.8	1685	2	T02750	acetyl-CoA carboxy
24	48	34.0	367	2	D97096	probable permease,
25	48	34.0	384	2	T06814	GTP-binding regula
26	48	34.0	454	2	AA6532	Ig mu chain C regi
27	48	34.0	958	2	S15566	pol protein - simi
28	48	34.0	1161	2	S18738	pol protein - simi
29	47	33.3	110	2	S43147	Ig upsilon chain -

30	47	33.3	320	2	C86148	hypothetical prote
31	47	33.3	352	2	AE0595	molybdenum transpo
32	47	33.3	372	2	T16971	probable GTP-bind1
33	47	33.3	384	1	RGTOOA	GTP-binding regula
34	47	33.3	2254	2	D86215	protein T6D22.14 [
35	47	33.3	2257	2	T09538	acetyl-CoA carboxy
36	46.5	33.0	242	1	C97419	probable phosphoe
37	46.5	33.0	242	2	AD2637	conserved hypotnet
38	46.5	33.0	248	2	T48376	hypothetical prote
39	46	32.6	198	2	B71044	hypothetical prote
40	46	32.6	245	2	I55951	MHC class II E-bet
41	46	32.6	264	2	A60497	H-2 class II histo
42	46	32.6	264	2	S10989	class II histocomp
43	46	32.6	382	2	T34940	probable membrane
44	46	32.6	549	2	S04845	Ig heavy chain pre
45	45.5	32.3	270	1	A64367	pyruvate synthase

ALIGNMENTS

RESULT 1

I36948
Ig epsilon-chain - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C:Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orang
A:Reference number: I36948; MUID:87147196
A:Accession: I36948
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-426 <RES>
A:Cross-references: GB:M15398; NID:gl76797; PIDN:AAA35416.1; PID:gl76798
C:Genetics:
A:Introns: 103/1; 209/1; 317/1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 62.4%; Score 88; DB 2; Length 426;
Best Local Similarity 69.6%; Pred No. 4.7e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 GETYYSRVTHPLPKDIVRSIAK 24
||||| |||||||: |||||
DB 292 GETYQCRVTHPLPRALVRSTTK 314
RESULT 2
EHHU
Ig epsilon chain C region - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1981 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116;
R:Flanagan, J.G.; Rabbitts, T.H.
EMBO J. 1, 655-660, 1982
A:Title: The sequence of a human immunoglobulin epsilon heavy chain constant region g
A:Reference number: A22771; MUID:84236029
A:Accession: A22771
A:Molecule type: DNA
A:Residues: 1-428 <FLA>
A:Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:gl85035
R:Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.
EMBO J. 1, 1539-1544, 1982
A:Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseu
A:Reference number: A23195; MUID:84207910
A:Accession: A23195
A:Molecule type: DNA
A:Residues: 2-428 <UED>
A:Cross-references: GB:J00222; NID:gl84755

Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 SRVTHPLPKDIVRSIAKC 25
II I IIIII:: II
Db 15 SRCVHHLPKDIVAIRPKC 33

RESULT 9

T19571
hypothetical protein C29F7.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T19571
R:McMurray, A.
submitted to the EMBL Data Library, March 1997
A:Reference number: Z19144
A:Accession: T19571
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-394 <WIL>
A:Cross-references: EMBL:Z92827; PIDN:CAB07326.1; GSPDB:GN00028; CESP:C29F7.1
A:Experimental source: clone C29F7
C:Genetics:
A:Gene: CESP:C29F7.1
A:Map position: X
A:Introns: 32/3; 72/3; 117/1; 164/3; 193/3; 247/1; 335/3; 363/3

Query Match 38.3%; Score 54; DB 2; Length 394;
Best Local Similarity 60.0%; Pred. No. 4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPHLPKDIVRSIAKC 25
II:IIII::II II
Db 62 HPNLPKNVVIKIASC 76

RESULT 10

S42659
acetyl-CoA carboxylase (EC 6.4.1.2) - maize (fragment)
C:Species: Zea mays (maize)
C:Date: 07-Sep-1994 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999
C:Accession: S42659; S34636
R:Ashton, A.R.; Jenkins, C.L.D.; Whitfield, P.R.
Plant Mol. Biol. 24, 35-49, 1994
A:Title: Molecular cloning of two different cDNAs for maize acetyl CoA carboxylase.
A:Reference number: S42659; MUID:94154242
A:Accession: S42659
A:Molecule type: mRNA
A:Residues: 1-1306 <ASH>
A:Cross-references: EMBL:Z24449
C:Function:
A:Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA
A:Pathway: fatty acid biosynthesis
C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C:Keywords: ligase

Query Match 38.3%; Score 54; DB 2; Length 1306;
Best Local Similarity 55.6%; Pred. No. 14;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
III II: III II::
Db 140 ETYISRLYQPHLVKDSIQ 157

RESULT 11

T02921
acetyl-CoA carboxylase (EC 6.4.1.2) (clone A3) - maize (fragment)
C:Species: Zea mays (maize)
C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C:Accession: T02921

R:Ashton, A.R.; Jenkins, C.L.D.; Whitfield, P.R.
Plant Mol. Biol. 24, 35-49, 1994
A:Title: Molecular cloning of two different cDNAs for maize acetyl CoA carboxylase.
A:Reference number: S42659; MUID:94154242
A:Accession: T02921
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1625 <ASH>
A:Cross-references: EMBL:Z24449; NID:q1272684; PIDN:CAA80822.1; PID:q1272685
A:Experimental source: strain B73; leaf
C:Function:
A:Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA
C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C:Keywords: ligase
F:68-140/Domain: lipoyl/biotin-binding homology <LPB>

Query Match 38.3%; Score 54; DB 2; Length 1625;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
III II: III II::
Db 459 ETYISRLYQPHLVKDSIQ 476

RESULT 12

T06161
acetyl-CoA carboxylase (EC 6.4.1.2) - wheat
N:Alternate names: acetyl-coenzyme A carboxylase
C:Species: Triticum aestivum (common wheat)
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 18-Aug-2000
C:Accession: T06161; T06162; S42660; S78600; S35959
R:Gornicki, P.; Faris, J.; Podkowinski, J.; Gill, B.; Haselkorn, R.
Proc. Natl. Acad. Sci. U.S.A. 94, 14179-14184, 1997
A:Title: Plastid localized acetyl-CoA carboxylase of bread wheat is encoded by a single
A:Reference number: Z15495; MUID:98054381
A:Accession: T06161
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-2311 <GOR1>
A:Cross-references: EMBL:AF029895; NID:q2827149; PIDN:AAC39330.1; PID:q2827150
A:Experimental source: cv. Hard Red Winter Tam 107, light-grown seedlings
A:Accession: T06162
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-10, 'F', 12-67, 'H', 69-502, 'N', 504-619 <GOR2>
A:Cross-references: EMBL:AF029897; NID:q2827153; PIDN:AAC39332.1; PID:q2827154
A:Experimental source: cv. Hard Red Winter Tam 107, light-grown seedling
R:Eiborough, K.M.; Simon, J.W.; Swinhoe, R.; Ashton, A.R.; Slabas, A.R.
Plant Mol. Biol. 24, 21-34, 1994
A:Title: Studies on wheat acetyl CoA carboxylase and the cloning of a partial cDNA.
A:Reference number: S42660; MUID:94154234
A:Accession: S42660
A:Molecule type: mRNA
A:Residues: 'WRIM', 1771-2026, 'K', 2028-2073, 'F', 2075-2108, 'V', 2110-2120, 'L', 2122, 'A',
A:Cross-references: EMBL:Z23038; NID:g396278; PIDN:CAA80573.1; PID:g396279
A:Accession: S78600
A:Molecule type: protein
A:Residues: 1847-1852, 'E', 1854-1863, 1947-1962, 'D', 1964; 2085-2091, 'HL', 2139-2155, 'I', 2
C:Genetics:
A:Gene: Acc-1
A:Introns: 69/1; 90/2; 209/3; 278/3; 332/3; 363/3; 395/3; 453/3; 485/1; 547/3; 564/3;
C:Function:
A:Description: catalyzes the ATP-dependent carboxylation of acetyl-CoA to malonyl-CoA
C:Superfamily: human acetyl-CoA carboxylase; biotin carboxylase homology; lipoyl/biotin-
C:Keywords: ligase
F:134-640/Domain: biotin carboxylase homology <BC>
F:767-839/Domain: lipoyl/biotin-binding homology <LPB>
F:806/Binding site: biotin (lys) (covalent) #status predicted

A:Gene: At2g14770; F26C24.9
A:Map position: 2
A:Introns: 42/1; 351/3; 418/1; 617/3; 659/3; 690/3; 725/1; 758/3; 778/3; 812/3; 902/3;

Query Match 36.9%; Score 52; DB 2; Length 1756;
Best Local Similarity 37.5%; Pred. No. 38;
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGEYYSRVTHPHLPKDIVRSIAK 24
| | :|:|:| | :| |
DB 1337 CRWTLLENWNPINPDDLLAKIAK 1360

RESULT 15
FPMS
N:Alternate names: AFP; alpha-1-fetoprotein precursor; alpha-fetoglobulin precursor
C:Species: Mus musculus (house mouse)
C:date: 01-Sep-1981 #sequence_revision 28-May-1986 #text_change 12-Jun-1998
C:Accession: A93254; A92305; A03235
R:Law, S.W.; Dugaiczky, A.
Nature 291, 201-205, 1981
A:title: Homology between the primary structure of alpha-fetoprotein, deduced from a
A:Reference number: A93254; MUID:81197641
A:Accession: A93254
A:Molecule type: mRNA
A:Residues: 1-597,'E',599-605 <LAW>
A:Cross-references: GB:V00743
R:Minghetti, P.P.; Law, S.W.; Dugaiczky, A.
Mol. Biol. Evol. 2, 347-358, 1985
A:title: The rate of molecular evolution of alpha-fetoprotein approaches that of pseu
A:Reference number: A93055; MUID:88216123
A:Contents: annotation: revision to reference 598
A:Note: residue 598, reported in reference A93254 as Glu, should have been reported a
R:Gorin, M.B.; Cooper, D.L.; Eiferman, F.; van de Rijn, P.; Tilghman, S.M.
J. Biol. Chem. 256, 1954-1959, 1981
A:title: The evolution of alpha-fetoprotein and albumin: I. A comparison of the prima
A:Reference number: A92305; MUID:81117287
A:Accession: A92305
A:Molecule type: mRNA
A:Residues: 15-532, 'RAKL',538-605 <GOR>
A:Cross-references: GB:M16381
A:Note: The beginning of the mature protein was placed at residue 21
R:Eiferman, F.A.; Young, P.R.; Scott, R.W.; Tilghman, S.M.
Nature 294, 713-718, 1981
A:title: Intragenic amplification and divergence in the mouse alpha-fetoprotein gene.
A:Reference number: A93271; MUID:82103990
A:Contents: annotation; exon-intron junctions
C:Comment: Mouse AFP has two carbohydrate chains and was found to bind estrogens and
C:Genetics:
A:Map position: 5
A:Introns: 29/1; 46/2; 86/3; 157/2; 201/3; 234/2; 277/3; 349/2; 393/3; 426/2; 472/3;
C:Superfamily: serum albumin; serum albumin repeat homology
F:1-20/Domain: fetus; globulin; glycoprotein; metal binding; plasma
F:21-605/Product: signal sequence #status predicted <SIG>
F:21-605/Product: alpha-fetoprotein #status predicted <MPT>
F:25-198/Domain: serum albumin repeat homology <SA1>
F:217-390/Domain: serum albumin repeat homology <SA2>
F:409-588/Domain: serum albumin repeat homology <SA3>
F:22/Binding site: copper (His) #status predicted
F:95-110,109-120,144-189,188-197,220-266,265-273,285-299,380-389,412-458,457-
F:247,498/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.5%; Score 51.5; DB 1; Length 605;
Best Local Similarity 57.9%; Pred. No. 15;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 6 YSRVTHPHLPKDIVRSIAK 24
| | | | | | :| |
DB 354 YSR-THNPVSVILRIAK 371

Search completed: August 26, 2002, 10:04:24
Job time: 70 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:17:02 ; Search time 24.08 seconds
(without alignments)
40.199 Million cell updates/sec

Title: US-09-701-623c-6
Perfect score: 141
Sequence: 1 CGETYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	60.3	428	1 EPC_HUMAN	P01854 homo sapien
2	82	58.2	429	1 EPC_RAT	P01855 rattus norv
3	58	41.1	421	1 EPC_MOUSE	P06336 mus musculu
4	54	38.3	116	1 YK1A_YEAST	P36078 saccharomyc
5	51.5	36.5	605	1 FETA_MOUSE	P02772 mus musculu
6	51	36.2	106	1 Y116_ADE02	P03287 human adeno
7	49.5	35.1	381	1 GVPC_HALM1	Q02228 halobacteri
8	49.5	35.1	382	1 GVPC1_HALM1	P24574 halobacteri
9	48	34.0	384	1 GBA2_PEA	O04279 pisum sativ
10	48	34.0	1161	1 POL_SFV1	P23074 simian foam
11	47	33.3	384	1 GBA1_LYCES	P26981 lycopersico
12	47	33.3	777	1 BAR1_HUMAN	Q99728 homo sapien
13	46	32.6	264	1 HB2D_RAT	P18211 rattus norv
14	46	32.6	670	1 UVRC_PSEFL	P32966 pseudomonas
15	45.5	32.3	270	1 KORB_METJA	Q57957 methanococc
16	45	31.9	242	1 VMAT_INCUJ	P12446 influenza c
17	45	31.9	319	1 CDK4_XENLA	Q91727 xenopus lae
18	45	31.9	626	1 R3C1_RICPR	Q92cx9 rickettsia
19	44.5	31.6	1235	1 IRS1_RAT	P35570 rattus norv
20	44	31.2	285	1 F2D8_XENLA	Q93274 xenopus lae
21	44	31.2	285	1 GVC2_HALN2	O48310 halobacteri
22	44	31.2	374	1 SPOP_HUMAN	O43791 homo sapien
23	43.5	30.9	450	1 MUC_CANFA	P01874 canis famil
24	43.5	30.9	530	1 CG2B_DROME	P20439 drosophila
25	43.5	30.9	581	1 F2D8_XENLA	Q93274 xenopus lae
26	43	30.5	130	1 ALB1_PEA	P08687 pisum sativ
27	43	30.5	260	1 CAH1_MOUSE	P13634 mus musculu
28	43	30.5	275	1 Y760_METJA	Q58170 methanococc
29	43	30.5	376	1 TPS2_MOUSE	O88856 mus musculu
30	43	30.5	391	1 MUCB_HUMAN	P04220 homo sapien
31	43	30.5	418	1 YMI3_YEAST	Q04526 saccharomyc
32	43	30.5	454	1 MUC_HUMAN	P01871 homo sapien
33	43	30.5	515	1 G6PD_TREPA	O83491 treponema p

ALIGNMENTS

RESULT 1

ID	EPC_HUMAN	STANDARD;	PRT;	428 AA.
AC	P01854;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig epsilon chain C region.			
GN	IGH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=83168897; PubMed=6300763;			
RA	Seno M., Kurakawa T., Ono Y., Onda H., Sasada R., Igarashi K.,			
RA	Kikuchi M., Sugino Y., Nishida Y., Honjo T.;			
RA	"Molecular cloning and nucleotide sequencing of human immunoglobulin			
RT	epsilon chain cDNA.";			
RL	Nucleic Acids Res. 11:719-726(1983).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=83001945; PubMed=6288268;			
RA	Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;			
RA	"Duplication and deletion in the human immunoglobulin epsilon genes.;"			
RL	Cell 29:691-699(1982).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=84236029; PubMed=62341164;			
RA	Flanagan J.G., Rabbitts T.H.;			
RA	"The sequence of a human immunoglobulin epsilon heavy chain constant			
RT	region gene, and evidence for three non-allelic genes.;"			
RL	EMBO J. 1:653-660(1982).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=84207910; PubMed=6327276;			
RA	Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;			
RA	"Long terminal repeat-like elements flank a human immunoglobulin			
RT	epsilon pseudogene that lacks introns.;"			
RL	EMBO J. 1:1539-1544(1982).			
RN	[5]			
RP	PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).			
RA	Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;			
RA	(In) Bach M.K. (eds.);			
RL	Immediate hypersensitivity: modern concepts and developments, pp.1-36,			
RL	Marcel Dekker, New York (1978).			
RN	[6]			
RP	SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.			
RA	MEDLINE=83065234; PubMed=6815656;			
RA	Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,			
RA	Bell L.O., Gould H.J.;			
RT	"Cloning and sequence determination of the gene for the human			
RT	immunoglobulin epsilon chain expressed in a myeloma cell line.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).			
RN	[7]			

34 43 30.5 1095 1 PIPA_DROME P13217 drosophila

35 43 30.5 1581 1 AR01_PNECA Q12659 p pentafunc

36 42.5 30.1 259 1 RFOD_PYRAB Q9V198 pyrococcus

37 42.5 30.1 332 1 YIE2_HSVB4 Q02484 bovine herp

38 42.5 30.1 550 1 SYM_CHLMU Q9P107 chlamydia m

39 42.5 30.1 611 1 FETA_RAT P02773 rattus norv

40 42.5 30.1 982 1 RPOL_KLULA P05472 kluyveromyc

41 42 29.8 248 1 YAUH_RHISN O53200 rhizobium s

42 42 29.8 264 1 YRBF_HAEIN P45031 haemophilus

43 42 29.8 275 1 TRUA_SYNY3 P73295 synechocyst

44 42 29.8 305 1 CDSA_CHLFR O84457 c phosphati

45 42 29.8 340 1 ALC2_HUMAN P01877 homo sapien

```

3D-STRUCTURE MODELING.
Padlan E.A., Davies D.R.;
"A model of the F-C of immunoglobulin E.";
Submitted (JUL-1993) to the PDB data bank.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; L00022; AAB59424.1; ALT_INIT.
DR PIR; A02142; EHHU.
DR PIR; A22771; A22771.
DR PIR; A23195; A23195.
DR PDB; 1IGE; 15-JUL-92.
DR MIM; 147180; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGC1; 4.
DR PROSITE; PS00290; IG_MHC; 3.
Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
3D-structure.
NON_TER 1 1
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 105
FT DISULFID 29 85
FT DISULFID 121 121 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 135 193
FT DISULFID 209 209 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 239 299
FT DISULFID 345 405
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .).
FT VARIANT 359 359 W -> L (IN REF. 2, POSSIBLY DUE TO
POLYMORPHISM).
FT FTID=VAR_003885.
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 60.3%; Score 85; DB 1; Length 428;
Best Local Similarity 65.2%; Pred. No. 3.5e-05;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24
|||||
Db 294 GETYQCRVTHPLPALMRSTTK 316
|||||

RESULT 2
EPC_RAT
ID EPC_MOUSE STANDARD; PRT; 429 AA.
AC P01855;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RA STRAIN=LOU/C/WSL;
RC MEDLINE=83064537; PubMed=6292865;
RX Hellman L., Pettersson U., Engstrom A., Karlsson T., Bennich H.;
RA

*Structure and evolution of the heavy chain from rat immunoglobulin
E.*;
Nucleic Acids Res. 10:6041-6049(1982).
[2]
SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6820340;
RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
construction, identification, and DNA sequence.";
RL DNA 1:335-343(1982).
[3]
SEQUENCE OF 205-306 FROM N.A.
RP MEDLINE=82174576; PubMed=6803238;
RX Hellman L., Pettersson U., Bennich H.;
RT "Characterization and molecular cloning of the mRNA for the heavy
(epsilon) chain of rat immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; J00744; AAA41379.1; ALT_INIT.
DR PIR; A02143; EHRT.
DR HSSP; P01854; LIGE.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00410; IG_like; 3.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS00290; IG_MHC; 3.
Immunoglobulin domain; Immunoglobulin C region.
NON_TER 1 1
FT CONFLICT 168 168 R -> N (IN REF. 2).
FT CONFLICT 308 308 P -> L (IN REF. 2).
SQ SEQUENCE 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;

Query Match 58.2%; Score 82; DB 1; Length 429;
Best Local Similarity 69.6%; Pred. No. 9.6e-05;
Matches 16; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24
|||||
Db 284 GETYQCRVTHPHFPKPIVRSITK 306
|||||

RESULT 3
EPC_MOUSE
ID EPC_MOUSE STANDARD; PRT; 421 AA.
AC P06336; P01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236082; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
comparison with the human epsilon gene sequence.";
RL EMBO J. 1:1117-1123(1982).
[2]
RP REVISIONS.
```

```
RA Honjo T.;
RL Submitted (APR-1986) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=8311774; PubMed=6818553;
RT Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
chain cDNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01857; CAA25977.1; -
DR EMBL; X01857; CAA25978.1; -
DR PIR; A02145; EHMS.
DR HSP; P01854; IIGE.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00410; Ig_Like; 2.
DR SMART; SM00407; IG1; 2.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.
FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;

Query Match 41.1%; Score 58; DB 1; Length 421;
Best Local Similarity 56.5%; Pred. No. 0.34;
Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24
ID YK14_YEAST STANDARD; PRT; 116 AA.
AC P36078;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE Hypothetical 13.6 kDa protein in MDH1-YMA5 intergenic region.
GN YK1084W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
```

```
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RL Pohl T.M., Pohl F.M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z28084; CAA81922.1; -
DR PIR; S37909; S37909.
DR SGD; S0001567; YKL084W.
KW Hypothetical protein.
SQ SEQUENCE 116 AA; 13641 MW; 57464E5FD9A591F8 CRC64;

Query Match 38.3%; Score 54; DB 1; Length 116;
Best Local Similarity 52.6%; Pred. No. 0.34;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 SRVTHPLPKDIVRSIAK 25
ID FETA_MOUSE STANDARD; PRT; 605 AA.
AC P02772;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Alpha-fetoprotein precursor (Alpha-fetoglobulin) (Alpha-1-
DE fetoprotein).
DE AFP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81197641; PubMed=6164927;
RX Law S.W., Dugaiczky A.;
RX "Homology between the primary structure of alpha-fetoprotein, deduced
RX from a complete cDNA sequence, and serum albumin.";
RT Nature 291:201-205(1981).
RN [2]
RN REVISION TO 598.
RX MEDLINE=88216123; PubMed=2452956;
RX Minghetti P.P., Law S.W., Dugaiczky A.;
RT "The rate of molecular evolution of alpha-fetoprotein approaches that
RT of pseudogenes.";
RT Mol. Biol. Evol. 2:347-358(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82103990; PubMed=6172714;
RX Eiferman F.A., Young P.R., Scott R.W., Tilghman S.M.;
RT "Intragenic amplification and divergence in the mouse
RT alpha-fetoprotein gene.";
RL Nature 294:713-718(1981).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RX Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RX Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RX Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RX Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
```

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [5]
 RP SEQUENCE OF 15-605 FROM N.A.
 RX MEDLINE=81117287; PubMed=6161929;
 RA Gorin M.B., Cooper D.L., Eiferman F.A., van de Rijn P., Tilghman S.M.;
 RT "The evolution of alpha-fetoprotein and albumin. I. A comparison of
 RT the primary amino acid sequences of mammalian alpha-fetoprotein and
 RT albumin.";
 RL J. Biol. Chem. 256:1954-1959(1981).
 CC -1- FUNCTION: BINDS ESTROGENS, FATTY ACIDS AND METALS.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- TISSUE SPECIFICITY: PLASMA.
 CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -1- PTM: GLYCOSYLATED; HAS TWO CARBOHYDRATE CHAINS.
 CC -1- PTM: SULFATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ALB/APP/VDB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; V00743; CAA24118.1; -;
 DR EMBL; M16394; AAA37189.1; JOINED.
 DR EMBL; M16381; AAA37189.1; JOINED.
 DR EMBL; M16382; AAA37189.1; JOINED.
 DR EMBL; M16383; AAA37189.1; JOINED.
 DR EMBL; M16384; AAA37189.1; JOINED.
 DR EMBL; M16385; AAA37189.1; JOINED.
 DR EMBL; M16386; AAA37189.1; JOINED.
 DR EMBL; M16387; AAA37189.1; JOINED.
 DR EMBL; M16388; AAA37189.1; JOINED.
 DR EMBL; M16389; AAA37189.1; JOINED.
 DR EMBL; M16390; AAA37189.1; JOINED.
 DR EMBL; M16391; AAA37189.1; JOINED.
 DR EMBL; M16392; AAA37189.1; JOINED.
 DR EMBL; M16393; AAA37189.1; JOINED.
 DR EMBL; AK010934; BAB27278.1; -;
 DR PIR; A03235; FPM5.
 DR HSP; P02768; IBJ5.
 DR MGD; MGI:87951; Afp.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 3.
 DR PRINTS; PR00802; SERUMALBUMIN.
 DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 3.
 DR PROSITE; PS00212; ALBUMIN; 2.
 KW Glycoprotein; Sulfation; Albumin; Plasma; Embryo; Repeat;
 KW Metal-binding; Signal
 FT SIGNAL 1 18 PROBABLE.
 FT CHAIN 19 605 ALPHA-FETOPROTEIN.
 FT REPEAT 28 198 1.
 FT REPEAT 217 390 2.
 FT REPEAT 409 588 3.
 FT CARBOHYD 247 247 N-LINKED (GLCNAC...) (POTENTIAL).
 FT *

FT CARBOHYD 325 325 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 498 498 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DISULFID 95 110
 FT DISULFID 109 120
 FT DISULFID 144 189
 FT DISULFID 188 197
 FT DISULFID 220 266
 FT DISULFID 265 273
 FT DISULFID 285 299
 FT DISULFID 298 309
 FT DISULFID 380 389
 FT DISULFID 412 458
 FT DISULFID 457 468
 FT DISULFID 481 497
 FT DISULFID 496 507
 FT DISULFID 534 579
 FT DISULFID 578 587
 FT CONFLICT 533 536
 SQ SEQUENCE 605 AA; 67337 MW; CE09EF50D74619A CRC64;
 LCQA -> RAKL (IN REF. 5).

Query Match 36.5%; Score 51.5; DB 1; Length 605;
 Best Local Similarity 57.9%; Pred. No. 4.5;
 Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 6 YSRVTHPLPKDIVRSTAK 24
 ||| |||:|:| |
 Db 354 YSR-THPLPVSVILRIAK 371

RESULT 6
 Y116_ADE02 STANDARD; PRT; 106 AA.
 AC P03287;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DE Hypothetical 11.6 kDa early protein.
 OS Human adenovirus type 2.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10515;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83056843; PubMed=7142161;
 RA Gingeras T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,
 RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
 RT "Nucleotide sequences from the adenovirus-2 genome.";
 RL J. Biol. Chem. 257:13475-13491(1982).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83056844; PubMed=7142162;
 RA Alestrom P., Akusjlaervi G., Pettersson M., Pettersson U.;
 RT "DNA sequence analysis of the region encoding the terminal protein
 RT and the hypothetical N-gene product of adenovirus type 2.";
 RL J. Biol. Chem. 257:13492-13498(1982).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J01917; -; NOT_ANNOTATED_CDS.
 DR PIR; A03859; A03859.
 KW Hypothetical protein; Early protein.
 SQ SEQUENCE 106 AA; 11668 MW; 12692EF62B02E3A3 CRC64;

Query Match 36.2%; Score 51; DB 1; Length 106;
 Best Local Similarity 62.5%; Pred. No. 0.87;
 Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;


```

QY 6 YSRVTHPLPKDIVRS 21
Db 23 YSRATHRHSGKTVRS 38

RESULT 7
GVPC_HALME STANDARD; PRT; 381 AA.
ID GVPC228;
AC Q02228;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gas vesicle protein C.
GN GVPC.
OS Halobacterium mediterranei (Haloferax mediterranei).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloferax.
OX NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1411;
RX MEDLINE=93021102; PubMed=1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
CC -!- FUNCTION: MAY CONFER STABILITY TO THE GAS VESICLE MEMBRANES. GAS
CC VESICLES ARE SMALL, HOLLOW, GAS FILLED PROTEIN STRUCTURES THAT ARE
CC FOUND IN SEVERAL PROKARYOTIC PLANKTONIC MICROORGANISMS. THEY ALLOW
CC THE POSITIONING OF THE BACTERIA, AT THE FAVORABLE DEPTH FOR
CC GROWTH.
CC -!- SUBCELLULAR LOCATION: BINDS TO THE EXTERNAL SURFACE OF THE GAS
CC VESICLE MEMBRANE.
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64701; CAA45944.1; -
DR PIR; S28115; S28115.
KW Gas vesicle; Repeat.
FT DOMAIN 22 274 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 59 1.
FT REPEAT 60 84 2.
FT REPEAT 85 122 3.
FT REPEAT 123 160 4.
FT REPEAT 161 192 5.
FT REPEAT 193 232 6.
FT REPEAT 233 274 7.
FT DOMAIN 246 335 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 381 AA; 42653 MW; 9FB48199D0305921 CRC64;

Query Match 35.1%; Score 49.5; DB 1; Length 381;
Best Local Similarity 64.3%; Pred. No. 5.5;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 CGEYYSRVTHPL 14
Db 343 CGEYYQAITPHL 355

RESULT 8
GVCL_HALNI
ID GVCL_HALNI STANDARD; PRT; 382 AA.
AC P24574; O9H117;
DT 01-MAR-1992 (Rel. 21, Created)

```

```

DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gas vesicle protein C 1.
GN (GVPC1 OR GVPC OR VNG5032G) AND (GVPC2 OR VNG6031G).
OS Halobacterium sp. (strain NRC-1), and
OS Halobacterium sp. (strain NRC-817).
OC Plasmid pNRC100, Plasmid pNRC200, and Plasmid pHHL.
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091, 148370;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=91323716; PubMed=1864501;
RA Jones J.G., Young D.C., Dassarma S.;
RT "Structure and organization of the gas vesicle gene cluster on the
Halobacterium halobium plasmid pNRC100.";
RL Gene 102:117-122(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=99063795; PubMed=9847077;
RA Ng W.V., Ciufo S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,
RA Hall B., Loretz C., Seto J., Slagel J., Hood L., DasSarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC200;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebbardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-817; PLASMID=pHHL;
RX MEDLINE=92065812; PubMed=1956294;
RA Horne M., Englert C., Wimmer C., Pfeifer F.;
RT "A DNA region of 9 kbp contains all genes necessary for gas vesicle
synthesis in halophilic archaeobacteria.";
RL Mol. Microbiol. 5:1159-1174(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-817; PLASMID=pHHL;
RX MEDLINE=93021102; PubMed=1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
CC -!- FUNCTION: MAY CONFER STABILITY TO THE GAS VESICLE MEMBRANES. GAS
CC VESICLES ARE SMALL, HOLLOW, GAS FILLED PROTEIN STRUCTURES THAT ARE
CC FOUND IN SEVERAL PROKARYOTIC PLANKTONIC MICROORGANISMS. THEY ALLOW
CC THE POSITIONING OF THE BACTERIA, AT THE FAVORABLE DEPTH FOR
CC GROWTH.
CC -!- SUBCELLULAR LOCATION: BINDS TO THE EXTERNAL SURFACE OF THE GAS
CC VESICLE MEMBRANE.
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58557; AAA98197.1; -.
DR EMBL; AF016485; AAC82810.1; -.
DR EMBL; AE005142; AAG20727.1; -.
DR EMBL; X57161; CAA40451.1; -.
DR EMBL; X64729; CAA45981.1; -.
DR PIR; J01122; J01122.
DR PIR; S15483; S15483.
DR PIR; S28128; S28128.
KW Gas vesicle; Plasmid; Repeat; Complete proteome.
FT DOMAIN 22 284 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 60 1.
FT REPEAT 61 92 2.
FT REPEAT 93 130 3.
FT REPEAT 131 168 4.
FT REPEAT 169 200 5.
FT REPEAT 201 240 6.
FT REPEAT 241 284 7.
FT DOMAIN 254 336 ASP/GLU-RICH (ACIDIC).
FT SEQUENCE 382 AA; 42391 MW; 171DBEB4C0364F46 CRC64;

Query Match 35.1%; Score 49.5; DB 1; Length 382;
Best Local Similarity 64.3%; Pred. No. 5.5;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 CGETYYSRVTHPLH 14
   ||| || : |||
Db 344 CGEYYQATEPHL 356

RESULT 9
GBA2_PEA
ID GBA2_PEA STANDARD; PRT; 384 AA.
AC O04279;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Guanine nucleotide-binding protein alpha-2 subunit (GP-alpha-2).
GN GPA2 OR GA2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALASKA; TISSUE=Apical bud;
RA Marsh J.F., Kaufman L.S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 2 (G(S)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U97044; AAB57826.1; -.
DR HSSP; P10824; IAS3.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
DR PRINTS; PR00318; GPROTEIN.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; ADP-ribosylation.
```

```
FT NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 219 223 GTP (BY SIMILARITY).
FT NP_BIND 288 291 GTP (BY SIMILARITY).
FT MOD_RES 191 191 ADP-RIBOSYL[1] (BY ACTION OF CTX).
SQ SEQUENCE 384 AA; 44669 MW; 6B0963393FD20BA4 CRC64;

Query Match 34.0%; Score 48; DB 1; Length 384;
Best Local Similarity 50.0%; Pred. No. 9.2;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 7 SRVTHPLPKDIVRSI 22
   || : ||| || : |
Db 125 SRLDYPHLTKDLAKEI 140

RESULT 10
POL_SFV1
ID POL_SFV1 STANDARD; PRT; 1161 AA.
AC P23074;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE POL polyprotein [Contains: Protease (EC 3.4.23.-); Reverse
DE transcriptase (EC 2.7.7.49); Endonuclease].
GN POL.
OS Simian foamy virus (type 1) (SFV-1).
OC Viruses; Retroid viruses; Retroviridae; Spumavirus.
OX NCBI_TaxID=11964;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91276270; PubMed=1647358;
RA Kupiec J.-J., Kay A., Hayat M., Ravier R., Perles J., Galibert F.;
RT "Sequence analysis of the simian foamy virus type 1 genome.";
RL Gene 101:185-194(1991).
RN [2]
RP SEQUENCE OF 1-970 FROM N.A.
RX MEDLINE=91361538; PubMed=1653483;
RA Mergia A., Luciw P.A.;
RT "Replication and regulation of primate foamy viruses.";
RL Virology 184:475-482(1991).
RN [3]
RP SEQUENCE OF 969-1161 FROM N.A.
RX MEDLINE=90080148; PubMed=2152825;
RA Mergia A., Shaw K.E.S., Lackner J.E., Luciw P.A.;
RT "Relationship of the env genes and the endonuclease domain of the pol
RT genes of simian foamy virus type 1 and human foamy virus.";
RL J. Virol. 64:406-410(1990).
CC -1- PTM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A9.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54482; -. NOT_ANNOTATED_CDS.
DR EMBL; X58484; CAA41394.1; -.
DR EMBL; M33561; AAA47793.1; -.
DR PIR; A33562; A33562.
DR PIR; S15566; S15566.
DR PIR; S18738; S18738.
DR HSSP; P03355; IMWL.
DR MEROPS; A09.001; -.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR000477; RVTse.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR001641; Spuma_A9ptase.
DR Pfam; PF00075; rnaseH; 1.
DR Pfam; PF00665; rve; 1.
```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; U76638; AAB38316.1; ..
CC EMBL; AF038042; AAB99978.1; ..
CC EMBL; AF038034; AAB99978.1; JOINED.
CC EMBL; AF038035; AAB99978.1; JOINED.
CC EMBL; AF038036; AAB99978.1; JOINED.
CC EMBL; AF038037; AAB99978.1; JOINED.
CC EMBL; AF038038; AAB99978.1; JOINED.
CC EMBL; AF038039; AAB99978.1; JOINED.
CC EMBL; AF038040; AAB99978.1; JOINED.
CC EMBL; AF038041; AAB99978.1; JOINED.
CC HSP; P15919; IRND.
CC MIM; 601593; -.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR001357; BRCT.
CC InterPro; IPR001841; Znf_ring.
CC Pfam; PF00023; ank; 3.
CC Pfam; PF00533; BRCT; 2.
CC PRINTS; PR01415; ANKYRIN.
CC SMART; SM00248; ANK; 3.
CC SMART; SM00184; RING; 1.
CC PROSITE; PS00088; ANK_REPEAT; 3.
CC PROSITE; PS00297; ANK_REPEAT; 1.
CC PROSITE; PS00172; BRCT; 2.
CC PROSITE; PS00518; ZF_RING_1; 1.
CC PROSITE; PS00089; ZF_RING_2; 1.
CC Zinc-finger; Nuclear protein; Disease mutation; Polymorphism; Repeat;
CC ANK repeat.
CC ZN_FING 50 87 RING-TYPE.
CC REPEAT 427 459 ANK 1.
CC REPEAT 460 492 ANK 2.
CC REPEAT 493 525 ANK 3.
CC DOMAIN 570 653 BRCT 1.
CC DOMAIN 667 777 BRCT 2.
CC DOMAIN 26 119 REQUIRED FOR INTERACTION WITH BRCA1.
CC DOMAIN 400 403 POLY-SER.
CC DOMAIN 542 545 POLY-LEU.
CC VARIANT 24 24
CC VARIANT 153 153 CAUCASIANS; LESS FREQUENT IN AFRICANS).
CC VARIANT 507 507 /FTid=VAR_010354.
CC VARIANT 507 507 K -> E (UNCLASSIFIED).
CC VARIANT 557 557 /FTid=VAR_010355.
CC VARIANT 557 557 V -> M (UNCLASSIFIED).
CC VARIANT 564 564 C -> S (RARE POLYMORPHISM IN CAUCASIANS).
CC VARIANT 564 564 /FTid=VAR_010357.
CC VARIANT 568 658 Q -> H (IN OVARIAN CANCER).
CC VARIANT 695 695 /FTid=VAR_010358.
CC VARIANT 695 695 R -> C (RARE POLYMORPHISM IN CAUCASIANS;
ABSENT IN AFRICANS).
CC VARIANT 761 761 /FTid=VAR_010359.
CC VARIANT 761 761 V -> L (IN BREAST AND OVARIAN CANCER).
CC VARIANT 761 761 /FTid=VAR_010360.
CC VARIANT 761 761 S -> N (IN UTERINE CANCER).
CC CONFLICT 378 378 /FTid=VAR_010361.
CC CONFLICT 406 406 R -> S (IN REF. 2).
CC CONFLICT 406 406 Q -> R (IN REF. 2).
CC SEQUENCE 777 AA; 86619 MW; AAE3C604524E802B CRC64;

Query Match 33.3%; Score 47; DB 1; Length 777;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 SRVTHPLPKDIVRSIAKC 25
| | | | | | | | | |
Db 602 STVTHVVPDVGAVQSLK 620

RESULT 13
HB2D_RAT
ID HB2D_RAT STANDARD; PRT; 264 AA.
AC P18211;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE RT1 class II histocompatibility antigen, D-1 beta chain precursor.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS FAMILIARIS; TISSUE=Bone marrow;
RX MEDLINE=90356406; PubMed=2388838;
RA Syha-Jedelhauser J., Reske K.;
RT "Sequence of rat cDNA clone plr beta 112 coding for the RT1.D beta I
chain.";
RL Nucleic Acids Res. 18:4598-4598(1990).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
THE IMMUNE SYSTEM.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```

```
CC EMBL; X53054; CAA37221.1; -.
CC PIR; S10989; S10989.
CC HSP; P13760; 2SEB.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003597; Ig_C1.
CC InterPro; IPR000353; MHC_II_beta.
CC Pfam; PF00047; Ig; 1.
CC Pfam; PF00969; MHC_II_beta; 1.
CC ProDom; PD000328; MHC_II_beta; 1.
CC SMART; SM00407; IG1; 1.
CC PROSITE; PS00290; IG_MHC; 1.
CC MHC II; Transmembrane; Glycoprotein; Signal.
CC SIGNAL 1 26
CC CHAIN 27 264 RT1 CLASS II HISTOCOMPATIBILITY ANTIGEN,
D-1 BETA CHAIN.
CC DOMAIN 27 120 EXTRACELLULAR BETA-1.
CC DOMAIN 121 215 EXTRACELLULAR BETA-2.
CC DOMAIN 216 226 CONNECTING PEPTIDE.
CC TRANSMEM 227 248
CC DOMAIN 249 264 CYTOPLASMIC TAIL.
CC DISULFID 42 106 BY SIMILARITY.
CC DISULFID 144 200 BY SIMILARITY.
CC CARBOHYD 46 46 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SEQUENCE 264 AA; 29876 MW; 158F357355177DA1 CRC64;

Query Match 32.6%; Score 46; DB 1; Length 264;
Best Local Similarity 47.1%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDI 18
| | | | | | | |
Db 195 GEVYTCQVHPSPV 211

RESULT 14
UVRC_PSEFL
ID UVRC_PSEFL STANDARD; PRT; 607 AA.
AC P32966;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
```

DE Excinuclease ABC subunit C.
GN UVRC.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL915;
RX MEDLINE=94355677; PubMed=8075420;
RA Gaffney T.D., Lam S.T., Ligon J., Gates K., Frazelle A., Maio J.,
RA Hill S., Goodwin S., Torkewitz N., Allshouse A.M., Kempf H.J.,
RA Becker J.O.;
RT "Global regulation of expression of antifungal factors by a
RT Pseudomonas fluorescens biological control strain.";
RL Mol. Plant Microbe Interact. 7:455-463(1994).
RN [2]
RP SEQUENCE OF 1-232 FROM N.A.
RC STRAIN=CHAO;
RX MEDLINE=92179223; PubMed=1311842;
RA Laville J., Voisard C.P., Keel C., Maurhofer M., Difago G.,
RA Haas D.;
RT "Global control in Pseudomonas fluorescens mediating antibiotic
RT synthesis and suppression of black root rot of tobacco.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1562-1566(1992).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE
CC UVRA-UVRB COMPLEX, REPLACING UVRA, AND THE DAMAGED DNA STRAND IS
CC NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UVRC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L29642; AAA98758.1; -
CC EMBL; M80913; AAA25822.1; -
CC InterPro; IPR000445; HHH.
CC InterPro; IPR003583; HHH_1.
CC InterPro; IPR001943; UVRC.
CC InterPro; IPR001162; UVRC_2.
CC InterPro; IPR000305; UvrC_1.
CC Pfam; PF01541; Exciendo_N; 1.
CC Pfam; PF00633; HHH; 2.
CC Pfam; PF02131; UVRC; 1.
CC ProDom; PD005870; UVRC_2; 1.
CC SMART; SM00465; GIYC; 1.
CC SMART; SM00278; HHH1; 2.
CC SOS response; Excision nuclease; DNA repair.
FT CONFLICT 3 4 DP -> EQ (IN REF. 2).
FT CONFLICT 28 28 T -> A (IN REF. 2).
FT CONFLICT 41 41 S -> N (IN REF. 2).
FT CONFLICT 44 44 A -> S (IN REF. 2).
FT CONFLICT 52 52 L -> Q (IN REF. 2).
FT CONFLICT 61 61 G -> A (IN REF. 2).
FT CONFLICT 111 111 E -> D (IN REF. 2).
FT CONFLICT 123 123 K -> R (IN REF. 2).
FT CONFLICT 184 184 A -> P (IN REF. 2).
FT CONFLICT 211 212 AG -> TA (IN REF. 2).
FT CONFLICT 215 215 Q -> A (IN REF. 2).
SQ SEQUENCE 607 AA; 67302 MW; 179D412EC0B4398C CRC64;

Query Match 32.6%; Score 46; DB 1; Length 607;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 GGETYSRVTHPLPKDIVRSIAKC 25
Db 154 CEDSYKNNRTRPCLOYQIKRKAPC 178
RESULT 15
KORB_METJA STANDARD; PRT; 270 AA.
AC Q57957;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 2-oxoglutarate synthase subunit KORB (EC 1.2.7.3) (2-ketoglutarate
DE oxidoreductase beta chain) (KOR) (2-oxoglutarate-ferredoxin
DE oxidoreductase beta subunit).
GN KORB OR MJ0537.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Keriavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:11058-1073(1996).
CC -1- CATALYTIC ACTIVITY: 2-OXOBUTANOATE + COA + OXIDIZED FERREDOXIN =
CC PYRUVYL-COA + CO(2) + REDUCED FERREDOXIN.
CC -1- SUBUNIT: HETEROTETRAMER OF THE KORA, KORB, KORC AND KORD
CC SUBUNITS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67503; AAB98531.1; -
CC TIGR; MJ0537; -
CC InterPro; IPR000399; TPP_enzyme.
CC Pfam; PF02775; TPP_enzymes_C; 1.
CC KOW Oxidoreductase; Complete proteome.
SQ SEQUENCE 270 AA; 29894 MW; 4F7331CF37216554 CRC64;
Query Match 32.3%; Score 45.5; DB 1; Length 270;
Best Local Similarity 52.2%; Pred. No. 15;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
QY 2 GGETYSRVTHPLPKDIVRSIAK 24
Db 164 GATYVARWTTAAH-PIQLVRSIKK 185
Search completed: August 26, 2002, 10:17:02
Job time: 698 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:16:17 ; Search time 66.82 Seconds
(without alignments)
64.724 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	54	38.3	394	5	017620 caenorhabdi
2	54	38.3	1079	10	Q94FR6 Q94fr6 zea mays (m
3	54	38.3	1079	10	Q94FR7 Q94fr7 zea mays (m
4	54	38.3	1625	10	Q08367 Q08367 zea mays (m
5	54	38.3	2311	10	Q48959 Q48959 triticum ae
6	54	38.3	2325	10	Q41743 Q41743 zea mays (m
7	52	36.9	1756	10	Q80975 Q80975 arabidopsis
8	51	36.2	195	17	Q9VIA1 Q9vial pyrococcus
9	51	36.2	416	2	Q9X4C5 Q9x4c5 escherichia
10	50	35.5	539	10	Q9LZX0 Q9lzx0 arabidopsis
11	49	34.8	405	11	Q91Z67 Q91z67 mus musculus
12	49	34.8	684	13	Q90544 Q90544 ginglymosto
13	49	34.8	1078	10	Q94FR5 Q94fr5 zea mays (m
14	49	34.8	1095	4	Q75044 Q75044 homo sapien
15	49	34.8	1685	10	Q43248 Q43248 zea mays (m
16	48	34.0	141	15	Q87052 Q87052 simian foam

17	48	34.0	367	16	Q97IP8 Q97ip8 clostridium
18	48	34.0	388	2	Q9KY27 Q9ky27 streptomyce
19	48	34.0	460	5	Q9V7W8 Q9v7w8 drosophila
20	47.5	33.7	421	5	Q9BIC9 Q9bic9 trichinella
21	47	33.3	320	10	Q9LQ87 Q9lq87 arabidopsis
22	47	33.3	372	10	Q40404 Q40404 nicotiana p
23	47	33.3	384	10	Q9FV62 Q9fv62 nicotiana t
24	47	33.3	518	13	Q9W6R2 Q9w6r2 fugu rubrip
25	47	33.3	542	3	Q9HFC4 Q9hfc4 zygosacchar
26	47	33.3	822	10	Q9FNM8 Q9fnm8 arabidopsis
27	47	33.3	2254	10	Q9LN02 Q9ln02 arabidopsis
28	47	33.3	2257	10	Q40326 Q40326 medicago sa
29	46.5	33.0	248	10	Q9LZD6 Q9lzd6 arabidopsis
30	46.5	33.0	550	5	Q9L145 Q9l145 lucilia cup
31	46	32.6	198	17	Q59359 Q59359 pyrococcus
32	46	32.6	245	7	Q31270 Q31270 rattus norv
33	46	32.6	261	7	Q9TQ47 Q9tq47 rattus norv
34	46	32.6	264	7	Q9TOA5 Q9tga5 rattus norv
35	46	32.6	382	2	Q9S2Y1 Q9s2y1 streptomyce
36	46	32.6	492	10	Q09484 Q09484 chlamydomon
37	46	32.6	1010	10	Q9ZPN1 Q9zpn1 avena sativ
38	46	32.6	1071	10	Q94FS1 Q94fs1 lolium rigi
39	46	32.6	1071	10	Q94FS0 Q94fs0 lolium rigi
40	46	32.6	1071	10	Q94FR9 Q94fr9 lolium rigi
41	46	32.6	1071	10	Q94FR8 Q94fr8 lolium rigi
42	46	32.6	1205	10	Q9ZNV6 Q9zav6 oryza sativ
43	46	32.6	1251	10	Q9SSY0 Q9ssy0 oryza sativ
44	45.5	32.3	213	2	Q9FCZ9 Q9fcz9 erwinia ste
45	45.5	32.3	342	5	Q9ULB1 Q9ulb1 leishmania

ALIGNMENTS

RESULT 1

017620 ID 017620 PRELIMINARY; PRT; 394 AA.
AC 017620;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE C29F7.1 PROTEIN.
GN C29F7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z92827; CAB07326.1; -.
DR InterPro; IPR004119; DUF227.
DR Pfam; PF02958; DUF227.1.
SQ SEQUENCE 394 AA; 44297 MW; D442AA972AE24451 CRC64;

Query Match 38.3%; Score 54; DB 5; Length 394;
Best Local Similarity 60.0%; Pred. No. 4;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPNLPKDIVRSIAKC 25
||:||||:| | |
DB 62 HPNLPKNNVIRKIASC 76

```
RESULT 2
Q94FR6 PRELIMINARY; PRT; 1078 AA.
AC Q94FR6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACETYL-COA CARBOXYLASE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, DK592;
RC MEDLINE=21287275; PubMed=11381131;
RA Zagnitko O., Jelenka J., Tevzadze G., Haselkorn R., Gornicki P.;
RT "An isoleucine/leucine residue in the carboxyltransferase domain of
RT acetyl-CoA carboxylase is critical for interaction with
RT arylloxyphenoxypionate and cyclohexanedione inhibitors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:6617-6622(2001).
DR EMBL; AF359518; AAK62314.1; -.
FT NON_TER 1
FT NON_TER 1078 1078
SQ SEQUENCE 1078 AA; 120645 MW; 0ACFC8BE8D794CD1 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1078;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
DB 350 ETYISRLYQPHLVKDSIQ 367

RESULT 3
Q94FR7 PRELIMINARY; PRT; 1079 AA.
AC Q94FR7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACETYL-COA CARBOXYLASE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, DK592;
RC MEDLINE=21287275; PubMed=11381131;
RA Zagnitko O., Jelenka J., Tevzadze G., Haselkorn R., Gornicki P.;
RT "An isoleucine/leucine residue in the carboxyltransferase domain of
RT acetyl-CoA carboxylase is critical for interaction with
RT arylloxyphenoxypionate and cyclohexanedione inhibitors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:6617-6622(2001).
DR EMBL; AF359517; AAK62313.1; -.
FT NON_TER 1
FT NON_TER 1079 1079
SQ SEQUENCE 1079 AA; 120717 MW; DAD7508EF98E9AE1 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1079;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
DB 350 ETYISRLYQPHLVKDSIQ 367
```

```
RESULT 4
Q08367 PRELIMINARY; PRT; 1625 AA.
AC Q08367;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACETYL-COA CARBOXYLASE (EC 6.4.1.2) (ACC)
DE [CONTAINS: BIOTIN CARBOXYLASE (EC 6.3.4.14)] (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B73; TISSUE=LEAF;
RC MEDLINE=94154242; PubMed=7906562;
RA Ashton A.R., Jenkins C.L., Whitfield P.R.;
RT "Molecular cloning of two different cDNAs for maize acetyl CoA
RT carboxylase.";
RL Plant Mol. Biol. 24:35-49(1994).
CC -!- FUNCTION: THIS PROTEIN CARRIES THREE CATALYTIC FUNCTIONS: BIOTIN
CC CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
CC CARBOXYLTRANSFERASE.
CC -!- CATALYTIC ACTIVITY: ATP + ACETYL-COA + HCO(3)(-) -> ADP +
CC ORTHOPHOSPHATE + MALONYL-COA.
CC -!- COFACTOR: BIOTIN.
CC -!- ENZYME REGULATION: BY PHOSPHORYLATION.
CC -!- PATHWAY: FIRST STEP (RATE LIMITING) IN LONG-CHAIN FATTY ACID
CC SYNTHESIS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: PARTIAL TO CARBAMYL PHOSPHATE SYNTHETASES.
DR EMBL; Z24449; CAA80822.1; -.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR Pfam; PF0364; biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme.
FT NON_TER 1
FT NON_TER 1625 1625
SQ SEQUENCE 1625 AA; 181392 MW; 3624EB042D1FF6B0 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1625;
Best Local Similarity 55.6%; Pred. No. 18;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
DB 459 ETYISRLYQPHLVKDSIQ 476

RESULT 5
Q48959 PRELIMINARY; PRT; 2311 AA.
AC Q48959;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ACETYL-COENZYME A CARBOXYLASE.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, HARD RED WINTER TAM 107;
RC MEDLINE=98054381; PubMed=9391173;
RA Gornicki P., Faris J., Podkowinski J., Gill B., Haselkorn R.;
```


RT "plastid-localized acetyl-CoA carboxylase of bread wheat is encoded by
a single gene on each of the three ancestral chromosome sets.";
Proc. Natl. Acad. Sci. U.S.A. 94:14179-14184 (1997).

RL EMBL; AF029895; AAC39330.1; -;
DR EMBL; AF029896; AAC39331.1; -;
DR HSSP; P24182; 1DV1.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR000901; CPSase.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 2311 AA; 254970 MW; 1BBC5CDD0F9F984B CRC64;

Query Match 38.3%; Score 54; DB 10; Length 2311;
Best Local Similarity 55.6%; Pred. No. 27;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
||| ||: ||| ||: |
Db 1159 ETYSRLYQPHLVKDSIQ 1176

RESULT 6
Q41743
ID Q41743 PRELIMINARY; PRT; 2325 AA.

AC Q41743;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ACETYL-COENZYME A CARBOXYLASE.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A188; TISSUE=LEAF;
RX MEDLINE=95357420; PubMed=7630949;
RA Egli M.A., Lutz S.M., Somers D.A., Gengenbach B.G.;
RT "A maize acetyl-coenzyme A carboxylase cDNA sequence.";
RL Plant Physiol. 108:1299-1300 (1995).
DR EMBL; U19183; AAA80214.1; -;
DR HSSP; P24182; 1DV1.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR000901; CPSase.
DR Pfam; PF02785; Biotin_carb_C; 1.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR Pfam; PF02786; CPSase_L_D2; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 2325 AA; 257150 MW; ECB0AEF3FE26FE0C CRC64;

Query Match 38.3%; Score 54; DB 10; Length 2325;
Best Local Similarity 55.6%; Pred. No. 27;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPLPKDIVR 20
||| ||: ||| ||: |
Db 1158 ETYSRLYQPHLVKDSIQ 1175

RESULT 7
O80975
ID O80975 PRELIMINARY; PRT; 1756 AA.

AC O80975;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AT2G14770 PROTEIN.
GN AT2G14770.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004705; AAC24188.1; -;
DR InterPro; IPR003653; SUMO_protease.
DR Pfam; PF02902; Ulpl_C; 1.
SQ SEQUENCE 1756 AA; 197368 MW; 68FC3B5FCC282775 CRC64;

Query Match 36.9%; Score 52; DB 10; Length 1756;
Best Local Similarity 37.5%; Pred. No. 40;
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAK 24
| | :|:|:| |:
Db 1337 CRNTLLNMTNPIDLLAKIAK 1360

RESULT 8
Q9V1A1
ID Q9V1A1 PRELIMINARY; PRT; 195 AA.
AC Q9V1A1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 21.8 KDA PROTEIN.
GN PAB0359.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248284; CAB49448.1; -;
DR InterPro; IPR000051; SAM_bind.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 195 AA; 21778 MW; 471178A62097C02C CRC64;

Query Match 36.2%; Score 51; DB 17; Length 195;
Best Local Similarity 47.8%; Pred. No. 5.5;
Matches 11; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

Qy	5	YRSRVTHPLPKDIVRSIAKC	25	
		I : : I I I I I : I		
Db	137	YFGLGLEHPLFPKDFIHDLIAC	157	
	RESULT	12		
	Q90544			
ID	Q90544	PRELIMINARY:	PRT;	684 AA.
AC	Q90544;			
DT	01-NOV-1996	(TrEMBLrel. 01, Created)		
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)		
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)		
DE	NOVEL ANTIGEN RECEPTOR PRECURSOR.			
OS	Ginglymostoma cirratum (Nurse shark).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;			
OC	Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;			
OC	Ginglymostomatidae; Ginglymostoma.			
OX	NCBI_TaxID=7801;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=SPLEEN;			
RX	MDLINE=95183140; PubMed=7877689;			
RA	Greenberg A.S., Avila D., Hughes M., Hughes E.C.,			
RA	Flaenjoik M.F.;			
RA	"A new antigen receptor gene family that undergoes rearrangement an			
RT	extensive somatic diversification in sharks."			
RL	Nature 374:168-173(1995).			
DR	EMBL: U18701; AAB48195.1; -.			
DR	HSSP; P01709; 2MCG.			
DR				

```

DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_6.
DR SMART: SM00407; Igc1; 4.
DR SMART: SM00410; Ig_like; 2.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_3.
KW Signal; Receptor.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 34.8%; Score 49; DB 13; Length 684;
Best Local Similarity 47.6%; Pred. No. 43;
Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 4 TYYSRVTHPLPKDIVRSIAK 24
   || | | | | | | | |
Db 642 TVSCLVGHPSLNRLIRSTNK 662

RESULT 13
Q94FR5
ID Q94FR5 PRELIMINARY; PRT; 1078 AA.
AC Q94FR5;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ACETYL-COA CARBOXYLASE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Lillipsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DK592;
RX MEDLINE=21287275; PubMed=11381131;
RA Zagnitko O., Jelenska J., Tevzadze G., Haselkorn R., Gornicki P.;
RT "An isoleucine/leucine residue in the carboxyltransferase domain of
RT acetyl-CoA carboxylase is critical for interaction with
RT arylloxyphenoxypionate and cyclohexanedione inhibitors.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:6617-6622(2001).
DR EMBL: AF3959519; AAK62315.1; -.
FT NON_TER 1 1
FT NON_TER 1078 1078
SQ SEQUENCE 1078 AA; 120560 MW; B76F56B4CC608E4E CRC64;

Query Match 34.8%; Score 49; DB 10; Length 1078;
Best Local Similarity 52.9%; Pred. No. 69;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 TYYSRVTHPLPKDIVR 20
   || | | | | | | | |
Db 351 TYISRLYQHLVKDSIQ 367

RESULT 14
I075044
ID I075044 PRELIMINARY; PRT; 1095 AA.
AC I075044;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE KIAA0456 PROTEIN (FRAGMENT).
GN KIAA0456.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98116662; PubMed=9455484;
RA Seki N., Ohira M., Nagase T., Ishikawa K., Miyajima N., Nakajima D.,
RA Nomura N., Ohara O.;
RT "Characterization of cDNA clones in size-fractionated cDNA libraries
RT from human brain.";
RL DNA Res. 4:345-349(1997).
DR EMBL: AB007925; BAA32301.1; -.
DR HSSP: Q60631; IGBO.
DR InterPro: IPR001060; FCH.
DR InterPro: IPR00198; RhoGAP.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00611; FCH; 1.
DR Pfam: PF00620; RhoGAP; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR SMART: SM00055; FCH; 1.
DR SMART: SM00324; RhoGAP; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
FT NON_TER 1 1
FT NON_TER 1095 1095
SQ SEQUENCE 1095 AA; 123752 MW; 27A8AD852B62B1A0 CRC64;

Query Match 34.8%; Score 49; DB 4; Length 1095;
Best Local Similarity 38.1%; Pred. No. 71;
Matches 8; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 5 YYSRVTHPLPKDIVRSIAK 25
   | : | | | | | | | |
Db 592 YFRGLEHPLFPKDFHDLMAC 612

RESULT 15
Q43248
ID Q43248 PRELIMINARY; PRT; 1685 AA.
AC Q43248;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ACETYL COA CARBOXYLASE (EC 6.4.1.2) (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Lillipsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B73; TISSUE=LEAF;
RX MEDLINE=94154242; PubMed=7906562;
RA Ashton A.R., Jenkins C.L., Whitfield P.R.;
RT "Molecular cloning of two different cDNAs for maize acetyl COA
RT carboxylase.";
RL Plant Mol. Biol. 24:35-49(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B73; TISSUE=LEAF;
RA Ashton A.R.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U58598; AAB01188.1; -.
DR InterPro: IPR000089; Biotin_lipoyl.
DR InterPro: IPR000022; Carboxyl_trans.
DR Pfam: PF00364; biotin_lipoyl; 1.
DR Pfam: PF01039; Carboxyl_trans; 1.
KW Ligase.
FT NON_TER 1 1
FT NON_TER 1685 1685
SQ SEQUENCE 1685 AA; 187594 MW; 51688B4B7CA7B116 CRC64;

Query Match 34.8%; Score 49; DB 10; Length 1685;
Best Local Similarity 52.9%; Pred. No. 11e+02;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```

Qy 4 TYYSRVTHPHLPKDIVR 20
|| ||: ||| ||: |:
Db 519 TYISRLYQPHLYKDSIQ 535

Search completed: August 26, 2002, 10:16:18
Job time: 709 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:07:01 ; Search time 82.75 Seconds
(without alignments)
33.557 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 CGETYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	100.0	25	AA199999	Optimised IgE-CH3
2	141	100.0	45	AA180019	IgE immunogenic pe
3	141	100.0	45	AA180083	IgE immunogenic pe
4	141	100.0	46	AA180020	IgE immunogenic pe
5	141	100.0	57	AA180081	IgE immunogenic pe
6	141	100.0	62	AA180080	IgE immunogenic pe
7	141	100.0	63	AA180084	IgE immunogenic pe
8	118	83.7	124	AAW24097	Partial canine imm
9	118	83.7	312	AA179995	Dog immunoglobulin
10	118	83.7	341	AA1806208	Immunogenic peptid
11	118	83.7	417	AAW23067	Canine IgE heavy c

12	118	83.7	426	17	AA197753	Canine IgE.
13	108	76.6	25	21	AA179998	Optimised IgE-CH3
14	108	76.6	25	21	AA191212	Modified human IgE
15	108	76.6	25	21	AA196802	Peptide sequence o
16	108	76.6	42	21	AA180014	IgE immunogenic pe
17	108	76.6	42	21	AA191215	Modified MVF Th ep
18	108	76.6	42	21	AA191216	Modified MVF Th ep
19	108	76.6	42	21	AA191217	Modified MVF Th ep
20	108	76.6	42	21	AA196804	Peptide sequence o
21	108	76.6	45	21	AA180007	IgE-CH3 domain ant
22	108	76.6	45	21	AA191218	Modified HBV surfa
23	108	76.6	45	21	AA196805	Peptide sequence o
24	108	76.6	46	21	AA180011	IgE immunogenic pe
25	108	76.6	46	21	AA191213	Modified MVF Th ep
26	108	76.6	46	21	AA191214	Peptide sequence o
27	108	76.6	46	21	AA196803	IgE immunogenic pe
28	108	76.6	56	21	AA180016	IgE immunogenic pe
29	108	76.6	59	21	AA180010	IgE immunogenic pe
30	108	76.6	60	21	AA180013	IgE immunogenic pe
31	108	76.6	60	21	AA180015	IgE immunogenic pe
32	108	76.6	63	21	AA180008	IgE-CH3 domain ant
33	108	76.6	63	21	AA180012	IgE immunogenic pe
34	108	76.6	63	21	AA191219	Inv epitope/modifi
35	108	76.6	63	21	AA196806	Peptide sequence o
36	105	74.5	25	21	AA180000	Optimised IgE-CH3
37	105	74.5	25	21	AA180077	Optimised IgE-CH3
38	101	71.6	60	21	AA180078	IgE immunogenic pe
39	95	67.4	20	18	AAW24102	Canine immunoglobi
40	91	64.5	345	21	AA1806207	Immunogenic peptid
41	86	61.0	17	21	AA150893	Antibody 15A.2 bin
42	86	61.0	17	21	AA150894	Antibody 15A.2 can
43	86	61.0	110	14	AA183324	Variant IgE - muta
44	85	60.3	106	20	AA1842620	Human IgE Fcpsi
45	85	60.3	110	14	AA1833304	IgE Fc epsilon 3.

ALIGNMENTS

RESULT 1
AA179999
ID AA179999 standard; Peptide; 25 AA.
XX
AC AA179999;
XX
DT 15-MAY-2000 (first entry)
XX
DE Optimised IgE-CH3 domain antigen peptide for dog IgE.
XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
OS Canis sp.
OS Synthetic.
XX
XX WO9967293-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13959.
XX
XX 20-JUN-1998; 98US-0100287.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY, Walfield AM;
XX
XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy

XX PS Claim 1; Page 99; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 141; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.2e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Db 1 CGETYYSRVTHPLPKDIVRSIAKC 25

|||||

RESULT 2

AAY80019

ID AAY80019 standard; Peptide; 45 AA.

XX AC AAY80019;

XX DT 15-MAY-2000 (first entry)

XX DE IgE immunogenic peptide conjugate SEQ ID NO:26.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy -

XX PS Claim 14; Page 76; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 45 AA;

Query Match 100.0%; Score 141; DB 21; Length 45;

Best Local Similarity 100.0%; Pred. No. 2.3e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Db 21 CGETYYSRVTHPLPKDIVRSIAKC 45

|||||

RESULT 3

AAY80083

ID AAY80083 standard; Peptide; 45 AA.

XX AC AAY80083;

XX DT 15-MAY-2000 (first entry)

XX DE IgE immunogenic peptide conjugate SEQ ID NO:90.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy -

XX PS Claim 14; Page 77; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity, and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX

SQ Sequence 45 AA;

Query Match 100.0%; Score 141; DB 21; Length 45;

Best Local Similarity 100.0%; Pred. No. 2.3e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Db 21 cgetyyrvthplpkdivrsiakc 45

RESULT 4

AAY80020

ID AAY80020 standard; Peptide; 46 AA.

XX AC AAY80020;

XX DT 15-MAY-2000 (first entry)

XX XX IgE immunogenic peptide conjugate SEQ ID NO:27.

XX XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

OS WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX and anti-asthmatic properties. (I) induces polyclonal antibodies

XX specific for a target effector site on the epsilon-heavy chain of IgE,

XX and so preventing triggering and activation of mast cells and basophils

XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX containing (I) are used for active immunisation against IgE-mediated

XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX dermatitis. Nucleic acids that encode these compounds are useful for

XX recombinant production of corresponding peptides or in DNA vaccines.

XX Conjugates of (I) that include a promiscuous T helper cell epitope

XX (functional in genetically diverse subjects), in addition to a B cell

XX target epitope, have increased immunogenicity and may include cyclic

XX constraints (disulfide bridge) to stabilise conformational features and

XX maximize cross-reactivity to the natural target. They induce safe

XX (non-anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino

XX acid sequences used in the exemplification of the present invention.

SQ Sequence 46 AA;

Query Match 100.0%; Score 141; DB 21; Length 46;

Best Local Similarity 100.0%; Pred. No. 2.4e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Db 22 cgetyyrvthplpkdivrsiakc 46

RESULT 5

AAY80081

ID AAY80081 standard; Peptide; 57 AA.

XX AC AAY80081;

XX XX 15-MAY-2000 (first entry)

XX XX IgE immunogenic peptide conjugate SEQ ID NO:88.

XX XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

OS WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX and anti-asthmatic properties. (I) induces polyclonal antibodies

XX specific for a target effector site on the epsilon-heavy chain of IgE,

XX and so preventing triggering and activation of mast cells and basophils

XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,

XX containing (I) are used for active immunisation against IgE-mediated

XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX dermatitis. Nucleic acids that encode these compounds are useful for

XX recombinant production of corresponding peptides or in DNA vaccines.

XX Conjugates of (I) that include a promiscuous T helper cell epitope

XX (functional in genetically diverse subjects), in addition to a B cell

XX target epitope, have increased immunogenicity and may include cyclic

XX constraints (disulfide bridge) to stabilise conformational features and

XX maximize cross-reactivity to the natural target. They induce safe

XX (non-anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino

XX acid sequences used in the exemplification of the present invention.

SQ Sequence 57 AA;

Query Match 100.0%; Score 141; DB 21; Length 57;

Best Local Similarity 100.0%; Pred. No. 3.1e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 DB 33 cgetyyysrvthplpkdivrsiakc 57

RESULT 6

AAV80080
 ID AAY80080 standard; Peptide; 62 AA.

XX AC AAY80080;

XX DT 15-MAY-2000 (first entry)

XX DE IgE immunogenic peptide conjugate SEQ ID NO:87.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 62 AA;

Query Match 100.0%; Score 141; DB 21; Length 62;
 Best Local Similarity 100.0%; Pred. No. 3.4e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 DB 38 cgetyyysrvthplpkdivrsiakc 62

RESULT 7

AAV80084

ID AAY80084 standard; Peptide; 63 AA.

XX AC AAY80084;

XX DT 15-MAY-2000 (first entry)

XX DE IgE immunogenic peptide conjugate SEQ ID NO:91.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 63 AA;

Query Match 100.0%; Score 141; DB 21; Length 63;
 Best Local Similarity 100.0%; Pred. No. 3.5e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 DB 39 cgetyyysrvthplpkdivrsiakc 63

RESULT 8

AAW24097
 ID AAW24097 standard; peptide; 124 AA.

XX AC AAW24097;

XX DT 21-NOV-1997 (first entry)

XX Partial canine immunoglobulin E protein.

XX Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
XX Canis familiaris.
XX JP09169795-A.
XX 30-JUN-1997.
XX 22-DEC-1995; 95JP-0334381.
XX 22-DEC-1995; 95JP-0334381.
XX (HITB) HITACHI CHEM CO LTD.
XX WPI; 1997-389423/36.
XX N-PSDB; AAT85646.
XX Canine immunoglobulin E peptide fragment and related DNA - useful
for the preparation of anti-canine immunoglobulin E antibody
XX Claim 1; Page 8; 12pp; Japanese.
XX This is a partial canine immunoglobulin E (IgE) protein. Peptide
fragments (AAW24098-106) containing at least five continuous amino acids
of this sequence are used for the preparation of anti-canine IgE
antibody. The anti-canine IgE antibody can be used for the diagnosis of
canine allergies.
XX Sequence 124 AA;
Query Match 83.7%; Score 118; DB 18; Length 124;
Best Local Similarity 95.7%; Pred. No. 2.5e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GETYSRVTHPLPKDIVRSIAK 24
DB 57 getyyrcvrtbphlpkdivrsiak 79
RESULT 9
AAW79995
ID AAY79995 standard; Protein; 312 AA.
XX AAY79995;
XX 15-MAY-2000 (first entry)
XX Dog immunoglobulin E epsilon heavy chain SEQ ID NO:2.
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX Immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX antibody; allergy; allergic disease; immunisation; anti-allergic;
XX anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX Canis sp.
XX WO9967293-A1.
XX 29-DEC-1999.
XX 21-JUN-1999; 99WO-US13959.
XX 20-JUN-1998; 98US-0100287.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY, Walfeld AM;
XX WPI; 2000-160578/14.
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy -
XX Example 1; Page 66-68; 155pp; English.
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
and anti-asthmatic properties. (I) induces polyclonal antibodies
specific for a target effector site on the epsilon-heavy chain of IgE,
and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
containing (I) are used for active immunisation against IgE-mediated
allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
dermatitis. Nucleic acids that encode these compounds are useful for
recombinant production of corresponding peptides or in DNA vaccines.
CC conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX Sequence 312 AA;
Query Match 83.7%; Score 118; DB 21; Length 312;
Best Local Similarity 95.7%; Pred. No. 7.4e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GETYSRVTHPLPKDIVRSIAK 24
DB 187 getyyrcvrtbphlpkdivrsiak 209
RESULT 10
AAB06208
ID AAB06208 standard; protein; 341 AA.
XX AAB06208;
XX 22-NOV-2000 (first entry)
XX Immunogenic peptide consisting of opossum CH2, dog CH3 and opossum CH4.
XX Dog; opossum: immunoglobulin E; IgE; vaccination; infection; allergy;
XX asthma; eczema; immunogenic peptide.
XX Chimeric - Didelphis virginiana.
XX Chimeric - Canis sp.
XX WO200025722-A2.
XX 11-MAY-2000.
XX 21-OCT-1999; 99WO-SE01896.
XX 02-NOV-1998; 98US-0106652.
XX 22-SEP-1999; 99US-0401636.
XX (RESI-) RESISTENTIA PHARM AB.
XX Hellman LT;
XX WPI; 2000-365342/31.
XX Immunogenic polypeptides useful for preventing the harmful effects of
immunoglobulin E in mammals -
XX Disclosure; Fig 2; 50pp; English.
XX The present sequence is an immunogenic peptide consisting of the heavy
chain constant regions 2 and 4 of the opossum IgE and the heavy chain
constant region 3 from the dog. It was shown to cause a stronger

CC polyclonal anti-self IgE response than peptides consisting of the same
 CC regions from one mammal. Immunogenic peptides, particularly those
 CC consisting of different heavy chain constant regions, can be used for
 CC vaccination in humans, against bacterial and viral infections and
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.
 XX
 SQ Sequence 341 AA;

Query Match 83.7%; Score 118; DB 21; Length 341;
 Best Local Similarity 95.7%; Pred. NO. 8.2e-10;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24
 ID AAW23067 standard; Protein; 417 AA.
 AC AAW23067;
 XX
 DT 19-FEB-1998 (first entry)
 XX
 DE Canine IgE heavy chain constant region (exon 1-4 product).
 XX
 KW IgE; immunoglobulin; antibody; heavy chain constant region;
 KW allergy; hypersensitivity; therapy; dog; antisense;
 KW immunomodulation.
 XX
 OS Canis familiaris.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 55 /note= "encoded by ACC"
 FT Misc-difference 56 /note= "encoded by TAC"
 FT Misc-difference 67 /note= "encoded by GCC"
 FT Misc-difference 83 /note= "encoded by NNT"
 FT Misc-difference 174 /note= "encoded by GGN"
 FT Misc-difference 175 /note= "encoded by NNG"
 FT Misc-difference 176 /note= "encoded by TGN"
 FT Misc-difference 203 /note= "encoded by TCC"
 FT Misc-difference 204 /note= "encoded by GAC"
 FT
 XX WO9730156-A2.
 PN
 XX
 PD 21-AUG-1997.
 XX
 PF 14-FEB-1997; 97WO-US02322.
 XX
 PR 14-FEB-1996; 96US-0601197.
 XX
 PA (IDEX-) IDEX LAB INC.
 PI
 XX Harris RA, Mermer B, Siefring AE;
 XX
 DR WPI: 1997-425031/39.
 DR N-PSDB; AAT79278.
 XX
 XX Isolated canine IgE heavy chain constant region DNA - useful to
 PT develop products for treatment of canine allergies and for
 PT immunomodulation in dogs
 XX

PS Disclosure; Page 35-39; 59pp; English.
 XX
 CC This polypeptide is encoded by exons 1-4 (see AAT79278) of canine
 CC IgE heavy chain constant region (epsilon) genomic DNA. Another
 CC polypeptide, comprising the exon 5 and 6 product, is given in
 CC AAW23068. Recombinant peptides encoded by exons 1-6 can be
 CC produced in eukaryotic or prokaryotic cells. Such peptides,
 CC and antibodies raised against them, are used in methods to treat
 CC the manifestation of allergy in dogs, e.g. to treatment Type I
 CC immediate hypersensitivity, and for immunomodulation.
 XX
 SQ Sequence 417 AA;

Query Match 83.7%; Score 118; DB 18; Length 417;
 Best Local Similarity 95.7%; Pred. NO. 1e-09;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24
 ID AAR97753 standard; Protein; 426 AA.
 AC AAR97753;
 XX
 DT 28-AUG-1996 (first entry)
 XX
 DE Canine IgE.
 XX
 KW IgE; immunoglobulin; antibody; vaccine; allergy; hypersensitivity.
 XX
 OS Canis familiaris.
 PN WO9614867-A1.
 XX
 PD 23-MAY-1996.
 XX
 PF 03-NOV-1995; 95WO-US13795.
 XX
 PR 09-NOV-1994; 94US-0336891.
 PR 09-NOV-1994; 94US-0336583.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Hollis GF, Patel MD;
 XX
 DR WPI: 1996-277321/28.
 DR N-PSDB; AAT29824.
 XX
 PT New DNA encoding canine IgE and IgA - useful in vaccines, antisense
 PT therapy, assays, drug screening, etc.
 XX
 PS Claim 11; Page 29-30; 49pp; English.
 XX
 CC The canine IgE amino acid sequence (AAR97753) was deduced from
 CC an isolated gene (AAT29824) obt'd. from a canine liver DNA library.
 CC The cloning of the IgE gene allows prodn. of large quantities of
 CC recombinant IgE using bacterial, yeast, mammalian, insect or
 CC viral systems. The IgE can be used in drug development (e.g.
 CC small molecule screening, assay development and anti-IgE
 CC antibody generation). Fragments of IgE can be used in vaccines
 CC or to prevent IgE-mediated hypersensitivity. The new sequence
 CC information permits targeted modulation of IgE-mediated immune
 CC responses.
 XX
 SQ Sequence 426 AA;

Query Match 83.7%; Score 118; DB 17; Length 426;

Best Local Similarity 95.7%; Pred. No. 1.1e-09; Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24
||||| ||||||| ||||||| |||
Db 289 getyyrvthplpkdivrsiak 311

RESULT 13
AAV79998
ID AAY79998 standard; Peptide; 25 AA.
XX
AC AAY79998;
XX
DT 15-MAY-2000 (first entry)
XX
DE Optimised IgE-CH3 domain antigen peptide for human IgE.
XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO9967293-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13959.
XX
PR 20-JUN-1998; 98US-0100287.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY, Walfield AM;
XX
DR WPI; 2000-160578/14.
XX
PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
XX
PS Claim 1; Page 21; 155pp; English.
XX
CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAV79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 1.3e-09;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAK 25
||||| ||||||| :||| ||
Db 1 cgetyyrvthplpkdivrsiak 25

RESULT 14
AAV91212
ID AAY91212 standard; peptide; 25 AA.
XX
AC AAY91212;
XX
DT 22-MAY-2000 (first entry)
XX
DE Modified human IgE CH3 domain, SEQ ID NO:92.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVP;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Example 6; Page 40; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a

CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and AAY91226-Y91227 are immunogens comprising a
 CC CEMP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CEMP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
 Best Local Similarity 72.0%; Pred. No. 1.3e-09;
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 ||||| ||||| ||||| :||| ||
 Db 1 cgetygsrvthphlpralmrsttkc 25

RESULT 15
 AAY68602
 ID AAY68602 standard; peptide; 25 AA.
 XX
 AC AAY68602;
 XX
 DT 05-MAY-2000 (first entry)
 XX
 DE Peptide sequence of the invention.
 XX
 KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX
 OS Unidentified.
 XX
 PN WO9966952-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13960.
 XX
 PR 20-JUN-1998; 98US-0100414.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 DR WPI; 2000-160562/14.
 XX
 XX
 PT New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer
 XX
 XX
 PS Disclosure; Page 92; 102pp; English.
 XX
 CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising

CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.
 XX
 SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
 Best Local Similarity 72.0%; Pred. No. 1.3e-09;
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 ||||| ||||| ||||| :||| ||
 Db 1 cgetygsrvthphlpralmrsttkc 25

Search completed: August 26, 2002, 10:07:01
 Job time: 227 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:05:18 ; Search time 34.18 Seconds
(without alignments)
17.865 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 CGETYSRVTHPHLPKDIVRSIAK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	118	83.7	426	1	US-08-336-583-2
2	118	83.7	426	5	PCT-US95-13795-2
3	108	76.6	25	3	US-09-100-414B-95
4	108	76.6	25	4	US-09-303-323-95
5	108	76.6	42	3	US-09-100-414B-98
6	108	76.6	42	3	US-09-100-414B-99
7	108	76.6	42	3	US-09-100-414B-100
8	108	76.6	42	4	US-09-303-323-98
9	108	76.6	42	4	US-09-303-323-99
10	108	76.6	42	4	US-09-303-323-100
11	108	76.6	45	3	US-09-100-414B-101
12	108	76.6	45	4	US-09-303-323-101
13	108	76.6	46	3	US-09-100-414B-96
14	108	76.6	46	3	US-09-100-414B-97
15	108	76.6	46	4	US-09-303-323-96
16	108	76.6	46	4	US-09-303-323-97
17	108	76.6	63	3	US-09-100-414B-102
18	108	76.6	63	4	US-09-303-323-102
19	85	60.3	106	2	US-08-232-539D-54
20	85	60.3	113	2	US-08-232-539D-56
21	74.5	52.8	119	2	US-08-464-025A-1
22	72	51.1	22	2	US-08-232-539D-19
23	72	51.1	24	2	US-08-232-539D-20
24	72	51.1	56	2	US-08-232-539D-18
25	67.5	47.9	109	4	US-08-466-163B-1
26	57	40.4	118	3	US-08-466-151-1
27	54	38.3	1313	2	US-08-244-537-2

28	54	38.3	2311	4	US-08-934-386-9	Sequence 9, Appli
29	54	38.3	2325	3	US-08-417-089-6	Sequence 6, Appli
30	54	38.3	2325	4	US-08-695-651-6	Sequence 6, Appli
31	54	38.3	2325	4	US-08-930-285-6	Sequence 6, Appli
32	54	38.3	2325	4	US-08-695-421-6	Sequence 6, Appli
33	46	32.6	331	2	US-08-646-981-17	Sequence 17, Appli
34	46	32.6	1205	4	US-09-330-330-1	Sequence 1, Appli
35	45	31.9	561	3	US-09-192-545-2	Sequence 1, Appli
36	45	31.9	2254	2	US-08-677-010-3	Sequence 3, Appli
37	45	31.9	2254	2	US-08-790-519-3	Sequence 3, Appli
38	44.5	31.6	1155	1	US-08-094-948A-29	Sequence 29, Appli
39	44.5	31.6	1155	5	PCT-US96-09319-29	Sequence 29, Appli
40	44	31.2	109	2	US-08-646-981-6	Sequence 6, Appli
41	43.5	30.9	43	2	US-08-488-161-31	Sequence 31, Appli
42	43.5	30.9	43	3	US-08-488-161-31	Sequence 31, Appli
43	43.5	30.9	43	5	PCT-US95-1193A-31	Sequence 31, Appli
44	43	30.5	106	1	US-08-399-106A-7	Sequence 7, Appli
45	43	30.5	106	1	US-08-433-105A-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-08-336-583-2
; Sequence 2, Application US/08336583
; Patent No. 5629415
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,583
; FILING DATE: 09-NOV-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-583-2

Query Match 83.7%; Score 118; DB 1; Length 426;
Best Local Similarity 95.7%; Pred. No. 3,1e-11;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 GETYSRVTHPHLPKDIVRSIAK 24
||||| |||||||||
DB 289 GETYICRVTHPHLPKDIVRSIAK 311

RESULT 2
PCT-US95-13795-2
; Sequence 2, Application PC/TUS9513795
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13795
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 192111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-13795-2

Query Match 83.7%; Score 118; DB 5; Length 426;
Best Local Similarity 95.7%; Pred. No. 3.1e-11;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24
Db 289 GETYYCRVTHPLPKDIVRSIAK 311

RESULT 3
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 76.6%; Score 108; DB 3; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAK 25
Db 1 CGETYYSRVTHPLPKDIVRSIAK 25

RESULT 4
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 76.6%; Score 108; DB 4; Length 25;


```
Best Local Similarity 72.0%; Pred. No. 4.6e-11; Indels 0; Gaps 0;
Matches 18; Conservative 3; Mismatches 4;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||
Db 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||

RESULT 5
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-98

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||
Db 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||

RESULT 6
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-99

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||
Db 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||

RESULT 7
US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-100
```

```
ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-99

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||
Db 1 CGETYYSRVTHPLPKDIVRSIAKC 25
      ||||| ||||| ||||| :||| ||

RESULT 7
US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-100
```

US-09-100-414B-100

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 9

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998

NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-751-6849
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 10

US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPHLPKDIVRSIAKC 25
||||| |||||||: :||| ||
Db 18 CGETYYSRVTHPHLPALMRSTTKC 42

RESULT 11
US-09-100-414B-101
Sequence 101, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-101

Query Match 76.6%; Score 108; DB 3; Length 45;
Best Local Similarity 72.0%; Pred. No. 9.2e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPHLPKDIVRSIAKC 25
||||| |||||||: :||| ||
Db 21 CGETYYSRVTHPHLPALMRSTTKC 45

RESULT 12
US-09-303-323-101
Sequence 101, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 76.6%; Score 108; DB 4; Length 45;
Best Local Similarity 72.0%; Pred. No. 9.2e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPHLPKDIVRSIAKC 25
||||| |||||||: :||| ||
Db 21 CGETYYSRVTHPHLPALMRSTTKC 45

RESULT 13
US-09-100-414B-96
Sequence 96, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-96

Query Match 76.6%; Score 108; DB 3; Length 46;
Best Local Similarity 72.0%; Pred. No. 9.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
DB 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 14
US-09-100-414B-97
Sequence 97, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 76.6%; Score 108; DB 3; Length 46;
Best Local Similarity 72.0%; Pred. No. 9.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
DB 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 15
US-09-303-323-96
Sequence 96, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-96

Query Match 76.6%; Score 108; DB 4; Length 46;
Best Local Similarity 72.0%; Pred. No. 9.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
DB 22 CGETYQSRVTHPLPALMRSTTKC 46

Search completed: August 26, 2002, 10:05:18
Job time: 124 sec

THIS PAGE BLANK (usPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:04:24 ; Search time 45.87 Seconds
(without alignments)
52.370 Million cell updates/sec

Title: US-09-701-623c-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPKPIVRSITK 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 71: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122	84.1	429	1 EHR7	Ig epsilon chain C
2	98	67.6	107	2 I68730	IgE chain C3 regio
3	98	67.6	107	2 I68726	IgE chain C3 regio
4	98	67.6	388	1 EHMS	Ig epsilon chain C
5	98	67.6	548	2 S38864	Ig epsilon chain C
6	85	58.6	423	1 EHMSS	Ig epsilon chain C
7	83	57.2	426	2 I36948	Ig epsilon chain C
8	80	55.2	428	1 EHHU	Ig epsilon chain C
9	51	35.2	245	2 I55951	MHC class II E-bet
10	51	35.2	264	2 A60497	H-2 class II histoc
11	51	35.2	264	2 S10989	class II histocomp
12	49	33.8	345	1 HLCHB4	MHC class I histoc
13	49	33.8	355	2 I51309	major histocompatl
14	49	33.8	355	2 T28149	MHC class I histoc
15	49	33.8	355	2 T28152	MHC class I histoc
16	49	33.8	430	2 AG0531	cell cycle protein
17	49	33.8	549	2 S04845	Ig heavy chain pre
18	49	33.8	1208	2 T23222	hypothetical prote
19	48.5	33.4	567	2 T33400	protein kinase C h
20	48.5	33.4	597	2 T33399	protein kinase C h
21	48.5	33.4	704	1 S60117	protein kinase C (
22	48	33.1	93	2 I54421	MHC RT1.B-beta2 -
23	48	33.1	110	2 S43147	Ig epsilon chain -
24	48	33.1	152	2 S14236	Ig gamma-1 chain C
25	48	33.1	237	2 C60497	H-2 class II histoc
26	48	33.1	266	2 A39260	MHC class II histo
27	48	33.1	266	2 B39260	MHC class II histo
28	48	33.1	324	1 GLMS	Ig gamma-1 chain C
29	48	33.1	393	1 GLMSM	Ig gamma-1 chain C

30	48	33.1	444	2 PC4436	monoclonal antibod
31	47.5	32.8	1198	2 T49726	hypothetical prote
32	47	32.4	322	2 PS0019	Ig gamma-2a chain
33	47	32.4	1413	2 T26467	hypothetical prote
34	47	32.4	2254	2 D86215	protein T6022.14 f
35	46.5	32.1	232	2 AG2525	hypothetical prote
36	46.5	32.1	302	2 E86267	hypothetical prote
37	46.5	32.1	510	2 E97307	probable cardiolip
38	46	31.7	86	2 D84566	hypothetical prote
39	46	31.7	116	2 E64392	desulfoferrodoxin
40	46	31.7	238	2 A45544	BoLA-DQ beta-1 - b
41	46	31.7	260	2 I45938	MHC cell surface g
42	46	31.7	277	1 MNIH32	32K nonstructural
43	46	31.7	314	2 AB3610	multidrug resistanc
44	46	31.7	328	2 I47161	Ig gamma 3 chain c
45	46	31.7	328	2 I47158	Ig gamma 1 chain c

ALIGNMENTS

RESULT 1

EHRT

Ig epsilon chain C region - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
C:Accession: A93442; A90937; A02143
R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A:Reference number: A93442; MUID:83064537
A:Accession: A93442
A:Molecule type: mRNA
A:Residues: 1-429 <HEL>
R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction
A:Reference number: A90937; MUID:83182019
A:Contents: myeloma IR162
A:Accession: A90937
A:Molecule type: mRNA
A:Residues: 'N', 169-307, 'L', 309-342 <KIN>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:19-80/Domain: immunoglobulin homology <IM1>
F:118-186/Domain: immunoglobulin homology <IM2>
F:223-291/Domain: immunoglobulin homology <IM3>
F:327-398/Domain: immunoglobulin homology <IM4>
F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 84.1%; Score 122; DB 1; Length 429;
Best Local Similarity 95.7%; Pred. No. 4.1e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
|||||
Db 284 GEGYQSRVDHPHPKPIVRSITK 306

RESULT 2

I68730

IgE chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
R:Shinkai, Y.; Nakachi, H.; Honjo, T.; Okumura, K.
Immunogenetics 27, 288-292, 1988
A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic aci
A:Reference number: I54443; MUID:88152907

A;Residues: 1-358,'L',360-428 <MAX>
A;Cross-references: GB:J00222; NID:g184755
A;Note: this sequence difference may be due to polymorphism
R:Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.
A;Reference number: A94418
A;Accession: A94418
A;Molecule type: protein
A;Residues: 'GAWPL',6,'X',8-16,'B',18-43,'B',45-52,55-92,95-97,'B',99-121,'B',123,'L',
A;Experimental source: myeloma protein Nd
R:Kenten, J.H.; Mølgård, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A;Title: Cloning and sequence determination of the gene for the human immunoglobulin
A;Reference number: A93933; MUID:83065234
A;Accession: B93933
A;Molecule type: mRNA
A;Residues: 1-40;68-114;427-428 <KEN>
A;Cross-references: GB:L00022; NID:g185035
R:Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A;Title: Purification and characterization of a recombinant human IgE Fc-epsilon frag
A;Reference number: S02438; MUID:88083554
A;Accession: S02438
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 98-352 <IKE>
R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A;Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produ
A;Reference number: A53116; MUID:94103254
A;Accession: A53116
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 320-428 <ZH>
A;Experimental source: myeloma U266-derived cell line AF-10
A;Note: sequence extracted from NCBI backbone (NCBI:141701, NCBIP:141702)
R:Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A;Title: Characterization of four novel epsilon chain mRNA and a comparative analysis
A;Reference number: A46536; MUID:93122085
A;Accession: A46536
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 382-426 <HEL>
A;Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A;Experimental source: B cell myeloma U-266
A;Note: sequence extracted from NCBI backbone (NCBIP:125297)
A;Accession: D46536
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 382-391 <HE2>
A;Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A;Experimental source: B cell myeloma U-266
A;Note: sequence extracted from NCBI backbone (NCBIP:125299)
A;Accession: A46536
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 401-428 <HE3>
A;Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A;Experimental source: B cell myeloma U-266
A;Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics;
A;Gene: GDB:IGHE
A;Cross-references: GDB:l19335; OMIM:147180
A;Map position: 14q32.33-14q32.33
A;Introns: 1/1; 104/1; 211/1; 319/1
C;Complex: An immunoglobulin heterotetramer subunit consists of two identical light (C
C;Superfamily: Immunoglobulin C region; immunoglobulin homology
C;Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunogl
F;22-87/Domain: immunoglobulin homology <IM1>
F;128-195/Domain: immunoglobulin homology <IM2>
F;232-301/Domain: immunoglobulin homology <IM3>

RESULT 11
S10989

Db 272 GDKYQCRVEHASLPQP 287
 : | | | | | | | | | |

RESULT 13

major histocompatibility complex class I glycoprotein haplotype B21 - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 21-Jan-2000
 C:Accession: I51309
 R:Fullton, J.E.; Thacker, E.L.; Bacon, L.D.; Hunt, H.D.
 Eur. J. Immunol. 25, 2069-2076, 1995
 A:Title: Functional analysis of avian class I (BFIV) glycoproteins by epitope tagging and
 A:Reference number: I51309; MUID:95347411
 A:Accession: I51309
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: mRNA
 A:Residues: 1-355 <FUL>
 A:Cross-references: GB:S78682; NID:g1042200; PIDN:AAB34945.1; PID:g1042201
 C:Genetics:
 A:Gene: BFIV21
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:213-278/Domain: immunoglobulin homology <IMM>

Query Match 33.8%; Score 49; DB 2; Length 355;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17
 | : | | | | | | | | | |
 Db 271 GDKYQCRVEHASLPQP 286

RESULT 14

T28149
 MHC class I histocompatibility antigen B-F alpha chain 2 - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T28149
 R:Milne, S.; Kaufman, J.; Beck, S.
 submitted to the EMBL Data Library, May 1998
 A:Description: DNA sequencing and analysis of the chicken major histocompatibility compl
 A:Reference number: Z20475
 A:Accession: T28149
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-355 <MIL>
 A:Cross-references: EMBL:AL023516; PIDN:CAA18969.1
 A:Experimental source: clone cB12
 C:Genetics:
 A:Gene: BPa2
 A:Map position: 16
 A:Introns: 22/1; 110/1; 201/1; 292/1; 328/1; 339/1; 350/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 33.8%; Score 49; DB 2; Length 355;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17
 | : | | | | | | | | | |
 Db 271 GDKYQCRVEHASLPQP 286

RESULT 15

T28152
 MHC class I histocompatibility antigen B-F IV alpha chain precursor - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 C:Accession: T28152
 R:Milne, S.; Kaufman, J.; Beck, S.

submitted to the EMBL Data Library, May 1998
 A:Description: DNA sequencing and analysis of the chicken major histocompatibility co
 A:Reference number: Z20475
 A:Accession: T28152
 A:Status: preliminary; translated from GB/EMBL/DBBJ
 A:Molecule type: DNA
 A:Residues: 1-355 <MIL>
 A:Cross-references: EMBL:AL023516; PIDN:CAA18972.1
 A:Experimental source: clone cB12
 C:Genetics:
 A:Gene: BPa1
 A:Map position: 16
 A:Introns: 22/1; 110/1; 201/1; 292/1; 328/1; 339/1; 350/1
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match 33.8%; Score 49; DB 2; Length 355;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17
 | : | | | | | | | | | |
 Db 271 GDKYQCRVEHASLPQP 286

Search completed: August 26, 2002, 10:04:25
 Job time: 71 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:03:14 ; Search time 82.75 Seconds

(without alignments)

33.557 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

```

1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	25	AA199998	Optimised IgE-CH3
2	140	100.0	25	AA191212	Modified human IgE
3	140	100.0	25	AA198602	Peptide sequence o
4	140	100.0	42	AA1980014	IgE immunogenic pe
5	140	100.0	42	AA191215	Modified MVF Th ep
6	140	100.0	42	AA191216	Modified MVF Th ep
7	140	100.0	42	AA191217	Modified MVF Th ep
8	140	100.0	42	AA198604	Peptide sequence o
9	140	100.0	45	AA1980007	IgE-CH3 domain ant
10	140	100.0	45	AA191218	Modified HBV surfa
11	140	100.0	45	AA198605	Peptide sequence o

12	140	100.0	46	21	AA1980011	IgE immunogenic pe
13	140	100.0	46	21	AA191213	Modified MVF Th ep
14	140	100.0	46	21	AA191214	Modified MVF Th ep
15	140	100.0	46	21	AA198603	Peptide sequence o
16	140	100.0	56	21	AA1980016	IgE immunogenic pe
17	140	100.0	59	21	AA1980010	IgE immunogenic pe
18	140	100.0	60	21	AA1980013	IgE immunogenic pe
19	140	100.0	60	21	AA1980015	IgE immunogenic pe
20	140	100.0	63	21	AA1980008	IgE-CH3 domain ant
21	140	100.0	63	21	AA1980012	IgE immunogenic pe
22	140	100.0	63	21	AA191219	Inv epitope/modifi
23	140	100.0	63	21	AA198606	Peptide sequence o
24	133	95.0	60	21	AA1980078	IgE immunogenic pe
25	117	83.6	106	20	AA1942620	Human IgE Fc epsilo
26	117	83.6	110	14	AA193304	IgE Fc epsilon 3.
27	117	83.6	110	14	AA193314	Variant IgE - muta
28	117	83.6	110	14	AA193315	Variant IgE - muta
29	117	83.6	110	14	AA193318	Variant IgE - muta
30	117	83.6	110	14	AA193319	Variant IgE - muta
31	117	83.6	110	14	AA193320	Variant IgE - muta
32	117	83.6	110	14	AA193321	Variant IgE - muta
33	117	83.6	110	14	AA193322	Variant IgE - muta
34	117	83.6	110	14	AA193327	Variant IgE - muta
35	117	83.6	110	14	AA193328	Variant IgE - muta
36	117	83.6	110	14	AA193329	Variant IgE - muta
37	117	83.6	110	14	AA193330	Variant IgE - muta
38	117	83.6	110	14	AA193331	Variant IgE - muta
39	117	83.6	110	14	AA193202	Variant IgE - muta
40	117	83.6	110	14	AA193203	Variant IgE - muta
41	117	83.6	110	14	AA193204	Variant IgE - muta
42	117	83.6	110	14	AA193205	Variant IgE - muta
43	117	83.6	110	14	AA193206	Variant IgE - muta
44	117	83.6	110	14	AA193207	Variant IgE - muta
45	117	83.6	110	14	AA193208	Variant IgE - muta

ALIGNMENTS

RESULT 1

AA199998

ID AA199998 standard; Peptide; 25 AA.

XX AA199998;

AC AA199998;

XX 15-MAY-2000 (first entry)

DT Optimised IgE-CH3 domain antigen peptide for human IgE.

DE XX

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Homo sapiens.

OS Synthetic.

XX WO9967293-A1.

PN 29-DEC-1999.

PD 21-JUN-1999;

XX 99WO-US13959.

PF 20-JUN-1998;

XX 98US-0100287.

PR (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY, Walfield AM;

PI WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy -

```
XX PS Claim 1: Page 21; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies
XX CC specific for a target effector site on the epsilon-heavy chain of IgE,
XX CC and so preventing triggering and activation of mast cells and basophils
XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe
XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX CC acid sequences used in the exemplification of the present invention.
XX CC Sequence 25 AA;
XX CC
XX CC Query Match 100.0%; Score 140; DB 21; Length 25;
XX CC Best Local Similarity 100.0%; Pred. No. 7.4e-15;
XX CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX CC
XX CC QY 1 CGETYSQSRVTHPLPALMRSTTKC 25
XX CC |||||
XX CC Db 1 cgetygsrvthplpalmrsttkc 25
XX CC
XX CC RESULT 2
XX CC AAY91212
XX CC ID AAY91212 standard; peptide; 25 AA.
XX CC AC AAY91212;
XX CC DT 22-MAY-2000 (first entry)
XX CC DE Modified human IgE CH3 domain, SEQ ID NO:92.
XX CC
XX CC Promiscuous T-cell epitope; measles virus F protein; MVF;
XX CC hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX CC luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
XX CC somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
XX CC foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
XX CC Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
XX CC cholesteryl ester transport protein; anti-arteriosclerotic.
XX CC
XX CC Homo sapiens.
XX CC OS Synthetic.
XX CC PN WO9966957-A2.
XX CC XX 29-DEC-1999.
XX CC XX 21-JUN-1999; 99WO-US13975.
XX CC XX 20-JUN-1998; 98US-0100412.
XX CC XX (UNBI-) UNITED BIOMEDICAL INC.
XX CC PI Wang CY;
XX CC DR WPI; 2000-160564/14.
XX CC
XX CC New artificial T helper cell epitope and derived immunogens with target
XX CC PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX CC PT or human immune deficiency virus
XX CC PT
XX CC Example 6; Page 40; 129pp; English.
XX CC
```

The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes and peptide immunogens along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesteryl ester transport protein (CETP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration; for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AAY91121 represents a promiscuous T helper epitope from the measles virus F (MVF) protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVF Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AAY90212 is a modified version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a CETP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory invasive protein epitope from Yersinia species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

Query Match 100.0%; Score 140; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.4e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYSQSRVTHPLPALMRSTTKC 25
|||
Db 1 cgetygsrvthplpalmrsttkc 25

RESULT 3
AAY68602
ID AAY68602 standard; peptide; 25 AA.
XX
AC AAY68602;
XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide sequence of the invention.

XX KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX OS Unidentified.
 XX PN WO9966952-A1.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13960.
 XX PR 20-JUN-1998; 98US-0100414.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX DR WPI; 2000-160562/14.
 XX PT New peptide immunogen containing luteinising hormone-releasing hormone
 XX antigen site and helper T cell epitope, for e.g. contraception and
 XX treatment of cancer
 XX PS Disclosure; Page 92; 102pp; English.
 XX CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.
 XX SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 7.4e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTK 25
 ID AAY80014 standard; Peptide; 42 AA.
 XX AAY80014;
 XX AC AAY80014;
 XX DT 15-MAY-2000 (first entry)
 XX DE IgE immunogenic peptide conjugate SEQ ID NO:21.
 XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX OS Unidentified.

PN WO9967293-A1.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13959.
 XX PR 20-JUN-1998; 98US-0100287.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY; Walfield AM;
 XX DR WPI; 2000-160578/14.
 XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 XX for immunization against allergy
 XX PS Claim 14; Page 76; 155pp; English.
 XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 XX acid sequences used in the exemplification of the present invention.
 XX SQ Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;
 Best Local Similarity 100.0%; Pred. No. 1.3e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTK 25
 ID AAY91215 standard; peptide; 42 AA.
 XX AAY91215
 XX AC AAY91215;
 XX DT 22-MAY-2000 (first entry)
 XX DE Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:95.
 XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CFP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.
 XX OS Chimeric - Measles virus.
 XX OS Chimeric - Homo sapiens.
 XX PN WO9966957-A2.
 XX PD 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.
XX 20-JUN-1998; 98US-0100412.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CV;
XX WPI: 2000-160564/14.
XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Example 6; Page 98-99; 129pp; English.
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MV) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human Ige
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CETP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
XX invention.

SQ Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.3e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGEYQSRVTHPLRALMRSTTKC 25
Db 18 cgetyqsrvtphplralmrsttkc 42
RESULT 6
AAY91216
ID AAY91216 standard; peptide; 42 AA.
XX
AC AAY91216;
XX
DT 22-MAY-2000 (first entry)
XX
DE Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:96.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Chimeric - Measles virus.
OS Chimeric - Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
PI Wang CV;
XX
WPI: 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Example 6; Page 99; 129pp; English.
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MV) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human Ige
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CETP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
XX invention.

AC AAY68604;
DT 05-MAY-2000 (first entry)
XX Peptide sequence of the invention.
DE
XX
KW Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX
OS Unidentified.
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /label= Ile, Met, Leu
FT Misc-difference 2
FT /label= Ser, Thr
FT Misc-difference 5
FT /label= Lys, Arg
FT Misc-difference 6
FT /label= Gly, Thr
FT Misc-difference 10
FT /label= His, Thr
FT Misc-difference 11
FT /label= Lys, Arg
FT Misc-difference 12
FT /label= Ile, Met, Leu
FT Misc-difference 14
FT /label= Gly, Thr
FT Misc-difference 15
FT /label= Ile, Met, Val
XX
XX WO9966952-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13960.
XX
XX 20-JUN-1998; 98US-0100414.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY;
XX
XX WPI; 2000-160562/14.
XX
XX New peptide immunogen containing luteinising hormone-releasing hormone
XX antigen site and helper T cell epitope, for e.g. contraception and
XX treatment of cancer -
XX
XX Disclosure; Page 95; 102pp; English.
XX
XX The specification describes peptide immunogens comprising a
XX synthetic helper T cell (Th) epitope and a target antigen, luteinising
XX hormone-releasing hormone (LHRH). The peptide immunogens cause
XX induction of a specific immune response to LHRH which is involved in
XX regulation of spermatogenesis, ovulation, oestrus, sexual development
XX and secretion of sex hormones. Provision of a promiscuous T helper
XX epitope (which is functional in genetically diverse subjects) provides
XX optimum immunogenicity to the B cell epitopes of the target antigen and
XX thus high antibody titres against the target antigen. The peptide
XX immunogens of the invention are used to vaccinate against mammalian LHRH,
XX for use as (reversible) contraceptive; control of hormone-dependent
XX tumours (cancer of prostate or breast, also endometriosis); to prevent
XX boar taint (and improve meat quality) and for immunocastration. The
XX present sequence appears in the specification.

SQ Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;

Best Local Similarity 100.0%; Pred. No. 1.3e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGETYQSRVTHPLRALMRSTTKC 25
DB 18 cgetyqsrvtphplralmrsttkc 42
RESULT 9
AAY80007
ID AAY80007 standard; Peptide; 45 AA.
XX
XX AAY80007;
XX
XX 15-MAY-2000 (first entry)
XX
XX IgE-CH3 domain antigen peptide antigenic site 15b SEQ ID NO:14.
XX
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
XX Unidentified.
XX
XX WO9967293-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13959.
XX
XX 20-JUN-1998; 98US-0100287.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY, Walfield AM;
XX
XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Claim 14; Page 75; 155pp; English.
XX
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY9994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX
XX Sequence 45 AA;

Query Match 100.0%; Score 140; DB 21; Length 45;

Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLRALMRSTTKC 25

DB 21 cgetyqsrvtphplralmrsttkc 45

```

RESULT 10
AA91218
ID AA91218 standard; peptide: 45 AA.
XX
AC AA91218;
XX
DT 22-MAY-2000 (first entry)
XX
DE Modified HBV surface Ag/IgE CH3 domain, SEQ ID NO:98.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Chimeric - Hepatitis B virus.
OS Chimeric - Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI: 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX
PS Example 6; Page 100; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CPTP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AA91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AA91122-Y91142,
CC AA91126 and AA911245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AA91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AA91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AA91156-Y91196, AA91227 and AA91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AA91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AA91200 is somatostatin, and AA91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AA91208 is a
CC human CD4 CDR2-like domain antigenic site, and AA91209-Y90211 are MVH Th

```

```

CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AA90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AA90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AA91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AA91221-Y91222 comprise this peptide and a Th
CC epitope. AA91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AA91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AA91228-Y91231 represent
CC CPTP-derived peptides and AA91232-Y91241 are immunogens comprising a
CC CPTP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AA91247 and AA91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AA91248-Y91251 and
CC AA91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AA91198 and AA91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
SQ Sequence 45 AA;

```

```

Query Match 100.0%; Score 140; DB 21; Length 45;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGETYQSRVTHPLPRALMRSTTKC 25
Db 21 cgetygsrvthplpralmrsttkc 45

```

```

RESULT 11
AA91218
ID AA91218 standard; peptide: 45 AA.
XX
AC AA91218;
XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide sequence of the invention.
XX
KW Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX
OS Unidentified.
XX
PN WO9966952-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13960.
XX
PR 20-JUN-1998; 98US-0100414.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI: 2000-160562/14.
XX
PT New peptide immunogen containing luteinising hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer
XX
PS Disclosure; Page 95; 102pp; English.
XX
CC The specification describes peptide immunogens comprising a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The peptide immunogens cause

```

CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.

XX Sequence 45 AA;

Query Match 100.0%; Score 140; DB 21; Length 45;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGETYQSRVTHPLPALMRSTTKC 25
 |||||
 Db 21 cgetyqsrvtphlpralmrsttkc 45

RESULT 12
 AAY80011
 ID AAY80011 standard; Peptide; 46 AA.

XX AC AAY80011;

XX DT 15-MAY-2000 (first entry)

XX DE IgE immunogenic peptide conjugate SEQ ID NO:18.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 XX for immunization against allergy

XX PS Claim 14; Page 75; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 46 AA;

Query Match 100.0%; Score 140; DB 21; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGETYQSRVTHPLPALMRSTTKC 25
 |||||
 Db 22 cgetyqsrvtphlpralmrsttkc 46

RESULT 13
 AAY91213
 ID AAY91213 standard; peptide; 46 AA.

XX AC AAY91213;

XX DT 22-MAY-2000 (first entry)

XX DE Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:93.

XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX OS Chimeric - Measles virus.

XX OS Chimeric - Homo sapiens.

XX PN WO9966957-A2.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13975.

XX PR 20-JUN-1998; 98US-0100412.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX WPI; 2000-160564/14.

XX PT New artificial T helper cell epitope and derived immunogens with target
 XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
 XX or human immune deficiency virus

XX PS Example 6; Page 98; 129pp; English.

XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen

CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MV) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitopes/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human Ige
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitopes/Ige CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitopes which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

Sequence 46 AA:

Query Match 100.0%; Score 140; DB 21; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.5e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGEYQSRVTHPLRALMRSTTK 25
 |||||
 Db 22 cgetyqsrvtphlpralmrsttk 46

RESULT 14

AAY91214

ID AAY91214 standard; peptide; 46 AA.

XX AAY91214;

AC AAY91214;

XX 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/Ige CH3 domain, SEQ ID NO:94.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; Ige; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Homo sapiens.

XX W09966957-A2.

PN 29-DEC-1999.

XX

PF 21-JUN-1999; 99WO-US13975.

XX

PR 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI: 2000-160564/14.

DR New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus

XX Example 6; Page 98; 129pp; English.

CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration; for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MV) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitopes/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human Ige
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitopes/Ige CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitopes which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 46 AA;

Query Match 100.0%; Score 140; DB 21; Length 46;

Best Local Similarity 100.0%; Pred. No. 1.5e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CGETYQSRVTHPHLPALMRSTTKC 25
Db 22 cgetygsrvthphlpralmrsttkc 46

RESULT 15
AAAY68603
ID AAY68603 standard; peptide; 46 AA.
XX
AC AAY68603;
XX
DT
XX
DE Peptide sequence of the invention.
XX
KW Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX
OS Unidentified.
XX

FH Key Location/Qualifiers
FT Misc-difference 4 /label= Ser, Thr
FT Misc-difference 7 /label= Lys, Arg
FT Misc-difference 8 /label= Gly, Thr
FT Misc-difference 12 /label= His, Thr
FT Misc-difference 13 /label= Lys, Arg
FT Misc-difference 16 /label= Gly, Thr

XX WO9966952-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13960.
XX
XX 20-JUN-1998; 98US-0100414.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY;
XX
XX WPI; 2000-160562/14.
XX
XX New peptide immunogen containing luteinising hormone-releasing hormone
XX antigen site and helper T cell epitope, for e.g. contraception and
XX treatment of cancer -
XX
XX Disclosure: Page 93; 102pp; English.

CC The specification describes peptide immunogens comprising a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The peptide immunogens cause
CC induction of a specific immune response to LHRH which is involved in
CC regulation of spermatogenesis, ovulation, oestrus, sexual development
CC and secretion of sex hormones. Provision of a promiscuous T helper
CC epitope (which is functional in genetically diverse subjects) provides
CC optimum immunogenicity to the B cell epitopes of the target antigen and
CC thus high antibody titres against the target antigen. The peptide
CC immunogens of the invention are used to vaccinate against mammalian LHRH,
CC for use as (reversible) contraceptive; control of hormone-dependent
CC tumours (cancer of prostate or breast, also endometriosis); to prevent
CC boar taint (and improve meat quality) and for immunocastration. The
CC present sequence appears in the specification.

SQ Sequence 46 AA;
Query Match 100.0%; Score 140; DB 21; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.5e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CGETYQSRVTHPHLPALMRSTTKC 25
Db 22 cgetygsrvthphlpralmrsttkc 46
Search completed: August 26, 2002, 10:07:01
Job time: 227 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:03:14 ; Search time 45.87 Seconds
(without alignments)
52.370 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGEYQSRVTHPLPRALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	117	83.6	428	1 EHHU	Ig epsilon chain C
2	113	80.7	426	2 I36948	Ig epsilon chain -
3	80	57.1	429	1 EHRT	Ig epsilon chain C
4	56	40.0	107	2 I68730	IgE chain C3 regio
5	56	40.0	107	2 I68726	IgE chain C3 regio
6	56	40.0	388	1 EHWS	Ig epsilon chain C
7	56	40.0	548	2 S38864	Ig epsilon chain C
8	54	38.6	684	2 S60266	novel antigen rece
9	52	37.1	320	2 C86148	hypothetical prote
10	50	35.7	1597	2 S65053	genome polyprotein
11	50	35.7	1601	2 S48699	178K protein - tob
12	49.5	35.4	1235	1 S16948	insulin receptor s
13	48	34.3	338	2 B31194	hypothetical prote
14	48	34.3	1065	2 T32054	cellulose synthase
15	48	34.3	2254	2 D82115	protein T6D22.14 [
16	47.5	33.9	381	2 S28115	gas-vesicle protei
17	47.5	33.9	382	2 J01122	gas-vesicle protei
18	47.5	33.9	382	2 T08243	gas-vesicle operon
19	47	33.6	242	1 MFIVCJ	matrix protein M1
20	47	33.6	1940	1 S04090	myosin heavy chain
21	47	33.6	1940	1 A24922	myosin heavy chain
22	47	33.6	1940	2 A29320	myosin heavy chain
23	46.5	33.2	1231	2 S30185	insulin receptor s
24	46	32.9	116	2 S37909	hypothetical prote
25	46	32.9	247	2 A27547	trypsin (EC 3.4.21
26	46	32.9	430	2 AG0531	cell cycle protein
27	46	32.9	550	1 VGBE18	glycoprotein E - h
28	46	32.9	842	2 C83458	conserved hypothet
29	46	32.9	1019	2 A83613	conserved hypothet

30	45	32.1	93	2	I54421
31	45	32.1	106	2	A03859
32	45	32.1	210	2	S76316
33	45	32.1	245	2	I55951
34	45	32.1	264	2	A60497
35	45	32.1	264	2	S10989
36	45	32.1	343	2	T32334
37	45	32.1	343	2	T21355
38	45	32.1	676	2	H87906
39	45	32.1	1038	1	MWRBCB
40	45	32.1	1934	2	I48153
41	45	32.1	1935	1	A37102
42	45	32.1	1935	1	S06006
43	45	32.1	1935	2	A59286
44	45	32.1	1937	2	I38055
45	45	32.1	1938	1	JX0178

ALIGNMENTS

RESULT 1

EHHU

Ig epsilon chain C region - human

C.Species: Homo sapiens (man)

C.Date: 31-Mar

C.Accession: A

R:Flanagan, J.C

EMBO J. 1, 655

A>Title: The se

A.Reference num

A.Accession: A2

A.Molecule type

A.Residues: 1-4

A.Cross-referen

R:Ueda, S.; Naki

EMBO J. 1, 1539

A>Title: Long te

A.Reference num

A.Accession: A23

A.Molecule type: DNA

A.Residues: 2-428 <UED>

A.Cross-references: GB:J00222; NID:gl84755

R:Zhang, K.; Saxon, A.; Max, E.E.

J. Exp. Med. 176, 233-243, 1992

A>Title: Two unusual forms of human immunoglobulin E encoded by alternative RNA splic

A.Reference number: PH1214; MUID:92308839

A.Accession: PH1214

A.Molecule type: DNA

A.Residues: 320-428 <ZHA>

A.Cross-references: EMBL:X63693; GB:S38668; NID:g32987

R:Seno, M.; Kurokawa, T.; Ono, Y.; Onda, H.; Sasada, R.; Igarashi, K.; Kikuchi, M.; S

Nucleic Acids Res. 11, 719-726, 1983

A>Title: Molecular cloning and nucleotide sequencing of human immunoglobulin epsilon

A.Reference number: A93491; MUID:83168897

A.Accession: A93491

A.Molecule type: mRNA

A.Residues: 1-428 <SEN>

A.Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:gl85035

R:Max, E.E.; Battey, J.; Ney, R.; Kirsch, I.R.; Leder, P.

Cell 29, 691-699, 1982

A>Title: Duplication and deletion in the human immunoglobulin epsilon genes.

A.Reference number: A90824; MUID:83001945

A.Accession: A90824

A.Molecule type: DNA

A.Residues: 1-358, 'L', 360-428 <MAX>

A.Cross-references: GB:J00222; NID:gl84755

A>Note: this sequence difference may be due to polymorphism

R:Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.

in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.

A.Reference number: A94418

A.Accession: A94418

A.Molecule type: protein

33 #text_change 16-Jul-1999
124; A94418; B93933; S02438; A53116;

psilon heavy chain constant region g

35; NID:gl85035
Honjo, T.

a human immunoglobulin epsilon pseu

A:Residues: 'GAWTL',6,'X',8-16,'B',18-43,'B',45-52,55-92,95-97,'B',99-121,'B',123,'L',124
A:Experimental source: myeloma protein Nd
R:Kenton, J.H.; Molgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.; C
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A:Title: Cloning and sequence determination of the gene for the human immunoglobulin eps
A:Reference number: A93933; MUID:83065234
A:Accession: B93933
A:Molecule type: mRNA
A:Residues: 1-40:68-114:427-428 <KEN>
A:Cross-references: GB:L00022; NID:g185035
R:Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A:Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragment
A:Reference number: S02438; MUID:88083554
A:Accession: S02438
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 98-352 <IKE>
R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A:Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces
A:Reference number: A53116; MUID:94103254
A:Accession: A53116
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 320-428 <ZH2>
A:Experimental source: myeloma U266-derived cell line AF-10
A:Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBI:P:141702)
R:Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A:Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of
A:Reference number: A46536; MUID:93122085
A:Accession: C46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-426 <HEL>
A:Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBI:P:125297)
A:Accession: D46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBI:P:123483)
C:Genetics:
A:Gene: GDB:IGHE
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:22-87/Domain: immunoglobulin homology <IM1>
F:128-195/Domain: immunoglobulin homology <IM2>
F:232-301/Domain: immunoglobulin homology <IM3>
F:338-407/Domain: immunoglobulin homology <IM4>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:15-105,29-85,139-239,345-405/Disulfide bonds: #status predicted
F:21,49,99,146,252,275/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:121,209/Disulfide bonds: interchain (to heavy chain) #status predicted

Query Match

83.6%; Score 117; DB 1; Length 428;

Best Local Similarity 95.7%; Pred. No. 5.8e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 2 GETYOSRVTHPLPALMRSTTK 24
Db 294 GETYOCRVTHPLPALMRSTTK 316
RESULT 2
I36948
Ig epsilon-chain - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C:Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orang
A:Reference number: I36948; MUID:87147196
A:Accession: I36948
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-426 <RES>
A:Cross-references: GB:M15398; NID:g176797; PIDN:AAA35416.1; PID:g176798
C:Genetics:
A:Introns: 103/1; 209/1; 317/1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:336-405/Domain: immunoglobulin homology <IMW>
Query Match 80.7%; Score 113; DB 2; Length 426;
Best Local Similarity 91.3%; Pred. No. 2.3e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 2 GETYOSRVTHPLPALMRSTTK 24
Db 292 GETYOCRVTHPLPALVRSTTK 314
RESULT 3
EHRT
Ig epsilon chain C region - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
C:Accession: A93442; A90937; A02143
R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A:Reference number: A93442; MUID:83064537
A:Accession: A93442
A:Molecule type: mRNA
A:Residues: 1-429 <HEL>
A:Experimental source: strain LOU/c/Wsl, immunocytoma IR2
R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction
A:Reference number: A90937; MUID:83182019
A:Contents: myeloma IR162
A:Accession: A90937
A:Molecule type: mRNA
A:Residues: 'N',169-307,'L',309-342 <KIN>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (c
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:19-80/Domain: immunoglobulin homology <IM1>
F:118-186/Domain: immunoglobulin homology <IM2>
F:223-291/Domain: immunoglobulin homology <IM3>
F:327-398/Domain: immunoglobulin homology <IM4>
F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match

57.1%; Score 80; DB 1; Length 429;
Best Local Similarity 60.9%; Pred. No. 0.00024;

A;Accession: A02144
A;Molecule type: mRNA
A;Residues: 1-388 <IUI>
C;Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220
C;Complex: An immunoglobulin heterotrimer subunit consists of two identical light (L) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin
F:1-44/Domain: immunoglobulin homology (fragment) <IM1>
F:81-149/Domain: immunoglobulin homology <IM2>
F:186-254/Domain: immunoglobulin homology <IM3>
F:290-361/Domain: immunoglobulin homology <IM4>
F:10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 40.0%; Score 56; DB 1; Length 388;
Best Local Similarity 47.8%; Pred. No. 0.98;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPLPRALMRSTTK 24
I I I I I I I : : I I I I
DB 247 GYGQCIVDHPDFPKPIVRSITK 269

RESULT 7
S38864
Ig epsilon chain C region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C;Accession: S38864
R;Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A;Description: Combination of a defined specificity and desired isotype by cloning of
A;Reference number: S38864
A;Accession: S38864
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-548 <KIP>
C;Cross-references: EMBL:D27397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C;Superfamily: immunoglobulin C region; immunoglobulin homology
F:353-421/Domain: immunoglobulin homology <IMM>

Query Match 40.0%; Score 56; DB 2; Length 548;
Best Local Similarity 47.8%; Pred. No. 1.4;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPLPRALMRSTTK 24
I I I I I I I : : I I I I
DB 414 GYGQCIVDHPDFPKPIVRSITK 436

RESULT 8
S60266
novel antigen receptor precursor - nurse shark
C;Species: Ginglymostoma cirratum (nurse shark)
C;Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 21-Jul-2000
C;Accession: S60266
R;Greenberg, A.S.; Avila, D.; Hughes, M.; Hughes, A.; McKinney, E.C.; Flajnik, M.F.
Nature 374, 168-173, 1995
A;Title: A new antigen receptor gene family that undergoes rearrangement and extensive
A;Reference number: S60266; MUID:95183140
A;Accession: S60266
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-684 <GRE>
A;Cross-references: EMBL:U18701; NID:g699442; PIDN:AAB48195.1; PID:g699443

Query Match 38.6%; Score 54; DB 2; Length 684;
Best Local Similarity 57.1%; Pred. No. 3.5;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 TYQSRVTHPLPRALMRSTTK 24
|| | | | | | | | | |
Db 642 TYSLVGHPSLNRDLIRSTNK 662

RESULT 9

C86148
hypothetical protein TlnG.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C:Accession: C86148
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lurocs, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C86148
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <STO>
A:Cross-references: GB:AE005172; NID:g8671838; PIDN:AAF78401.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: acyl-CoA thiolesterase II

Query Match 37.1%; Score 52; DB 2; Length 320;
Best Local Similarity 60.0%; Pred. No. 3.2;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 ETYQSRVTHPLPR 17

|| | | | | | | | | |
Db 142 ELRSRITDPLPRS 156

RESULT 10

S65053
genome polyprotein - Chinese rape mosaic virus
N:Alternate names: RNA replicase
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: Chinese rape mosaic virus
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 17-Mar-1999
C:Accession: S65053
R:Aguiar, I.; Sanchez, F.; Martin Martin, A.; Martinez-Herrera, D.; Ponz, F.
Plant Mol. Biol. 30, 191-197, 1996
A:Title: Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic virus), a
A:Reference number: S65053; MUID:96197410
A:Accession: S65053
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-1597 <AGU>
A:Cross-references: EMBL:U30944
A:Note: readthrough of the terminator UGA occurs between codons CAA for 1103-Gln and CAA
A:Note: the internal stop codon is translated as X
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, July 1995
C:Superfamily: cucumber mosaic virus RNA 1 protein
C:Keywords: nucleotidyltransferase

Query Match 35.7%; Score 50; DB 2; Length 1597;
Best Local Similarity 36.8%; Pred. No. 35;
Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

QY 2 GETYQSRVTHPLPRALMRSTTK 25

|||||: ||| | | | | | | | | |
Db 1031 GETYKTAIVRLTATPLEIISRSPHVLVLRHTTRC 1068

RESULT 11

S48699
178K protein - tobacco mosaic virus (strain cr-TMV)
N:Alternate names: readthrough protein
N:Contains: 122K protein
C:Species: tobacco mosaic virus, TMV
A:Variety: strain cr-TMV
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 20-Sep-1999
C:Accession: S48699; S48659
R:Dorokhov, Y.L.; Ivanov, P.A.; Novikov, V.K.; Agranovsky, A.A.; Morozov, S.Y.; Efimo
FEBS Lett. 350, 5-8, 1994
A:Title: Complete nucleotide sequence and genome organization of a tobamovirus infect
A:Reference number: S48659; MUID:94341372
A:Accession: S48699
A:Molecule type: genomic RNA
A:Residues: 1-1601 <DOR>
A:Cross-references: EMBL:229370; NID:g488713; PIDN:CAA82559.1; PID:g619908
A:Experimental source: tobamovirus infecting cruciferae plants (cr-TMV)
A:Note: readthrough of the terminator UGA occurs between codons CAA for 1107-Gln and
A:Note: the internal stop codon is translated as X
C:Superfamily: cucumber mosaic virus RNA 1 protein
F:1-1601/Product: 178K protein #status predicted <PRO2>
F:1-1107/Product: 122K protein #status predicted <PRO1>

Query Match 35.7%; Score 50; DB 2; Length 1601;
Best Local Similarity 36.8%; Pred. No. 35;
Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

QY 2 GETYQSRVTHPLPRALMRSTTK 25

|||||: ||| | | | | | | | | |
Db 1035 GETYKTAIVRLTATPLEIISRSPHVLVLRHTTRC 1072

RESULT 12

S16948
Insulin receptor substrate IRS-1 - rat
N:Alternate names: insulin receptor substrate ppl85
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S16948; A39811
R:Sun, X.J.; Rothenberg, P.; Kahn, C.R.; Backer, J.M.; Araki, E.; Wilden, P.A.; Cahil
Nature 352, 73-77, 1991
A:Title: Structure of the insulin receptor substrate IRS-1 defines a unique signal tr
A:Reference number: S16948; MUID:91287824
A:Accession: S16948
A:Molecule type: mRNA
A:Residues: 1-1235 <SUN>
A:Cross-references: EMBL:X58375; NID:g56503; PIDN:CAA41264.1; PID:g56504
R:Rothenberg, P.L.; Lane, W.S.; Karasik, A.; Backer, J.; White, M.; Kahn, C.R.
J. Biol. Chem. 266, 8302-8311, 1991
A:Title: Purification and partial sequence analysis of ppl85, the major cellular subs
A:Reference number: A39811; MUID:91217066
A:Accession: A39811
A:Molecule type: protein
A:Residues: 44-51,173-178;223-243;489-506;635-646;871-882,'I',884,'X',886-888;932-936
A:Note: the phosphorylation residue was not identified
C:Comment: This protein and the beta chain of the insulin receptor itself are the maj
C:Comment: Phosphorylation of this protein in response to insulin is maximal at 30 se
insulin.
C:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology
C:Keywords: phosphoprotein; signal transduction
F:11-113/Domain: pleckstrin repeat homology <PLK>
F:872-891/Region: glutamine-rich

Query Match 35.4%; Score 49.5; DB 1; Length 1235;
Best Local Similarity 47.8%; Pred. No. 32;
Matches 11; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

QY 1 CGETYQSRVTHPLPR 16

|| | | | | | | | | |

```
Query Match      34.3%; Score 48; DB 2; Length 2254;
Best Local Similarity 26.3%; Pred. No. le-02;
Matches 5; Conservative 11; Mismatches 3; Indels 0; Caps 0;

QY    3 ETYSRVTHPHLPALMRMS 21
       :|:::|||||:::
Db     1985 KTFEEKVSFPHPKASILKA 2003
```

Search completed: August 26, 2002, 10:04:23
Job time: 69 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:05:24 ; Search time 24.08 seconds
(without alignments)
40.199 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETVQSRVTHPLRALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	83.6	428	1 EPC_HUMAN	P01854 homo sapien
2	80	57.1	429	1 EPC_RAT	P01855 rattus norv
3	56	40.0	421	1 EPC_MOUSE	P06336 mus musculu
4	50	35.7	1597	1 RRPO_CRMV	O66220 chinese rap
5	49.5	35.4	1235	1 IRS1_RAT	P35570 rattus norv
6	48	34.3	544	1 MKR3_MOUSE	O60764 mus musculu
7	47.5	33.9	381	1 GVPC_HALME	O02228 halobacteri
8	47.5	33.9	382	1 GVC1_HALN1	P24574 halobacteri
9	47	33.6	242	1 VMAI_INCIJ	P12446 influenza c
10	47	33.6	1938	1 MYHD_HUMAN	O9uxx3 homo sapien
11	47	33.6	1940	1 MYH3_CHICK	P02565 gallus gall
12	47	33.6	1940	1 MYH3_HUMAN	P11055 homo sapien
13	47	33.6	1940	1 MYH3_RAT	P12847 rattus norv
14	46.5	33.2	1233	1 IRS1_MOUSE	P35569 mus musculu
15	46	32.9	116	1 YK14_YEAST	P36078 saccharomyc
16	46	32.9	247	1 TRV3_RAT	P08426 rattus norv
17	46	32.9	550	1 VGLIE_HSV11	P04488 herpes simp
18	45	32.1	106	1 Y116_ADE02	P03287 human adeno
19	45	32.1	264	1 HB2D_RAT	P18211 rattus norv
20	45	32.1	736	1 MYH7_RABIT	P04461 oryctolagus
21	45	32.1	777	1 BAR1_HUMAN	O99728 homo sapien
22	45	32.1	1601	1 RRPO_TVCV	O88920 turnip vein
23	45	32.1	1934	1 MYH7_MESAU	P13540 mesocricetu
24	45	32.1	1935	1 MYH7_HUMAN	P12883 homo sapien
25	45	32.1	1935	1 MYH7_PIG	P79293 sus scrofa
26	45	32.1	1935	1 MYH7_RAT	P02564 rattus norv
27	45	32.1	1937	1 MYH8_HUMAN	P13335 homo sapien
28	45	32.1	1938	1 MYSS_CHICK	P13338 gallus gall
29	45	32.1	1939	1 MYH4_HUMAN	P12882 homo sapien
30	45	32.1	1939	1 MYH4_HUMAN	O9y623 homo sapien
31	45	32.1	1941	1 MYH2_HUMAN	O9uxx2 homo sapien
32	45	32.1	2492	1 TALA_DICDI	P54633 dictyosteli
33	44.5	31.8	450	1 MUC_CANFA	P01874 canis famil

ALIGNMENTS

RESULT 1

ID	EPC_HUMAN	STANDARD;	PRT;	428 AA.
AC	P01854;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ig epsilon chain C region.			
GN	IGH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83168897; PubMed=6300763;			
RA	Sano M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,			
RA	Kikuchi M., Sugino Y., Nishida Y., Honjo T.;			
RA	"Molecular cloning and nucleotide sequencing of human immunoglobulin			
RT	epsilon chain cDNA.";			
RL	Nucleic Acids Res. 11:719-726(1983).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=83001945; PubMed=6288268;			
RA	Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;			
RA	"Duplication and deletion in the human immunoglobulin epsilon genes.;"			
RL	Cell 29:691-699(1982).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84236029; PubMed=6234164;			
RA	Flanagan J.G., Rabbitts T.H.;			
RA	"The sequence of a human immunoglobulin epsilon heavy chain constant			
RT	region gene, and evidence for three non-allelic genes.;"			
RL	EMBO J. 1:655-660(1982).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=84207910; PubMed=6327276;			
RA	Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;			
RA	"Long terminal repeat-like elements flank a human immunoglobulin			
RT	epsilon pseudogene that lacks introns.;"			
RL	EMBO J. 1:1539-1544(1982).			
RN	[5]			
RP	PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).			
RA	Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;			
RL	(In) Bach M.K. (eds.);			
RL	Immediate hypersensitivity: modern concepts and developments, pp.1-36,			
RL	Marcel Dekker, New York (1978).			
RN	[6]			
RP	SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.			
RX	MEDLINE=83055234; PubMed=6815656;			
RA	Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,			
RA	Bell L.O., Gould H.J.;			
RA	"Cloning and sequence determination of the gene for the human			
RT	immunoglobulin epsilon chain expressed in a myeloma cell line.;"			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).			
RN	[7]			

34	44	31.4	340	1	ALC2_HUMAN	P01877 homo sapien
35	44	31.4	345	1	HA1F_CHICK	P15979 gallus gall
36	44	31.4	425	1	HE47_CAEEL	Q18212 caenorhabdi
37	44	31.4	427	1	HE47_PIG	Q29024 sus scrofa
38	44	31.4	428	1	HE47_HUMAN	Q13838 homo sapien
39	44	31.4	428	1	HE47_RAT	O63413 rattus norv
40	44	31.4	465	1	OPCA_NOSPU	P48971 nostoc punc
41	44	31.4	593	1	COX1_HALHA	P33518 halobacteri
42	44	31.4	1075	1	NFC3_HUMAN	Q12968 homo sapien
43	43.5	31.1	454	1	MUC_MESAU	P06337 mesocricetu
44	43.5	31.1	557	1	C791_SORBI	Q43135 sorghum bic
45	43	30.7	80	1	RL31_MYCTU	Q10608 mycobacteri

```

3D-STRUCTURE MODELING.
RA Padlan E.A., Davies D.R.;
RT "A model of the F-C- of immunoglobulin E.";
RL Submitted (JUL-1993) to the PDB data bank.
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L00022; AAB59424.1; ALT_INIT.
DR PIR; A02142; EHHU.
DR PIR; A22771; A22771.
DR PIR; A23195; A23195.
DR PDB; 1IGE; 15-JUL-92.
DR MIM; 147180; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00407; Igcl; 4.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW 3D-structure.
FT NON_TER 1 1
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 15 105
FT DISULFID 29 85
FT DISULFID 121 121
FT DISULFID 135 193
FT DISULFID 209 209 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 239 299
FT DISULFID 345 405 INTERCHAIN (WITH A HEAVY CHAIN).
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .).
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .).
FT VARIANT 359 W -> L (IN REF. 2, POSSIBLY DUE TO
FT POLYMORPHISM).
FT /FTIG-VAR_003885.
SQ SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 83.6%; Score 117; DB 1; Length 428;
Best Local Similarity 95.7%; Pred. No. 1.2e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHPLRALMRSTTK 24
DB 294 GETYQSRVTHPHPLRALMRSTTK 316
||||| ||||||| ||||||| |||||||
RESULT 2
EPC_RAT STANDARD; PRT; 429 AA.
AC P01555;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
RC STRAIN=LOU/C/WSL; PubMed=6292865;
RA HELLMAN U., Pettersson U., Engstroem A., Karlsson T., Bennich H.;

"Structure and evolution of the heavy chain from rat immunoglobulin
E.";
RL Nucleic Acids Res. 10:6041-6049(1982).
[2]
RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
RX MEDLINE=83182019; PubMed=6020340;
RA Kindvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
RT construction, identification, and DNA sequence.";
RL DNA 1:335-343(1982).
[3]
RP SEQUENCE OF 205-306 FROM N.A.
RX MEDLINE=82174576; PubMed=6803238;
RA Hellman L., Pettersson U., Bennich H.;
RT "Characterization and molecular cloning of the mRNA for the heavy
RT (epsilon) chain of rat immunoglobulin E.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; J00744; AAA41379.1; ALT_INIT.
DR PIR; A02143; EHRT.
DR HSP; P01854; IIGE.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00410; Ig-like; 3.
DR SMART; SM00407; Igcl; 1.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT CONFLICT 168 168 R -> N (IN REF. 2).
FT CONFLICT 308 308 P -> L (IN REF. 2).
FT CONFLICT 429 429 MW; D2970B34E8A72B0 CRC64;
SQ SEQUENCE 429 AA; 48671 MW;

Query Match 57.1%; Score 80; DB 1; Length 429;
Best Local Similarity 60.9%; Pred. No. 6.3e-05;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHPLRALMRSTTK 24
DB 284 GETYQSRVTHPHPLRALMRSTTK 306
||||| || ||| ||| ||| ||| |||
RESULT 3
EPC_MOUSE STANDARD; PRT; 421 AA.
AC P06336; P01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig epsilon chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=842336092; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT comparison with the human epsilon gene sequence.";
RL EMBO J. 1:1117-1123(1982).
RN [2]
RP REVISIONS.
```



```
RA Hoojo T.;
RN Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.
RX [3]
SEQUENCE OF 34-421 FROM N.A.
MEDLINE-8311774; PubMed=6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
chain cDNA."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01857; CAA25977.1; -
DR EMBL; X01857; CAA25978.1; -
DR PIR; A02145; EHMSS.
DR HSSP; P01854; LIQE.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00410; IG_Like; 2.
DR SMART; SM00407; IG_C1; 2.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 90
FT CH1.
FT DOMAIN 91 197
FT CH2.
FT DOMAIN 198 304
FT CH3.
FT DOMAIN 305 421
FT CH4.
FT DISULFID 23 75
FT BY SIMILARITY.
FT DISULFID 121 180
FT BY SIMILARITY.
FT DISULFID 226 285
FT BY SIMILARITY.
FT DISULFID 330 392
FT BY SIMILARITY.
FT CARBOHYD 43 43
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415
FT N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;

Query Match 40.0%; Score 56; DB 1; Length 421;
Best Local Similarity 47.8%; Pred. No. 0.31;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

Qy 2 GETYSRVTHPLPALMRSTTK 24
| | | | | | | | | | | |
Db 280 GYGQCIVDHPDFPKPIVRSITK 302

RESULT 4
RRPO_CRMV
ID RRPO_CRMV STANDARD; PRT; 1597 AA.
AC Q66220;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (182 kDa protein) [Contains:
DE Methyltransferase/RNA helicase (MT/HEL) (125 kDa protein)].
OC Chinese rape mosaic virus (CRMV) (Oilseed rape mosaic virus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=42007;
```

```
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-96197410; PubMed=8616237;
RA Aguilar I., Sanchez F., Martin-Martin A., Martinez-Herrera D.,
RA Ponz F.;
RT "Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic
virus), a crucifer tobamovirus infectious on Arabidopsis thaliana."
RL Plant Mol. Biol. 30:191-197(1996).
CC -!- FUNCTION: THE LONGER PROTEIN IS A RNA POLYMERASE ACTIVE IN VIRAL
CC RNA REPLICATION.
CC -!- FUNCTION: THE SMALLER PROTEIN IS A METHYLTRANSFERASE ACTIVE IN RNA
CC CAPPING AND AN RNA HELICASE.
CC -!- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UAG OCCURS
CC BETWEEN CODONS FOR GLN-1103 AND GLN-1105.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30944; AAB60599.1; -
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR002588; V_methyltransf.
DR InterPro; IPR000606; Viral_helicasel.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
DR Pfam; PF01443; Viral_helicasel; 1.
DR Pfam; PF01660; Vmethyltransf; 1.
KW Transferase; RNA-directed RNA polymerase; Helicase; ATP-binding.
FT CHAIN 1 1597
FT RNA-DIRECTED RNA POLYMERASE.
FT NP_BIND 823 830
FT METHYLTRANSFERASE/RNA HELICASE.
FT VARIANT 919 919
FT V -> G.
FT VARIANT 1286 1286
FT H -> Y.
SQ SEQUENCE 1597 AA; 181621 MW; 7256A908BD3308F2 CRC64;

Query Match 35.7%; Score 50; DB 1; Length 1597;
Best Local Similarity 36.8%; Pred. No. 11;
Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

Qy 2 GETYQS----RVT-----HPHPLPALMRSTTK 25
| | | | | | | | | | | |
Db 1031 GETYEKTAIVRLTATPLEIISRASPHVLVALTRHTTRC 1068

RESULT 5
IRSL_RAT
ID IRSL_RAT STANDARD; PRT; 1235 AA.
AC P35570;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Insulin receptor substrate-1.
GN IRS1 OR IRS-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE-91287824; PubMed=1648180;
RA Sun X.J., Rothenberg P., Kahn C.R., Backer J.M., Araki E.,
RA Wilden P.A., Cahill D.A., Goldstein B.J., White M.F.;
RT "Structure of the insulin receptor substrate IRS-1 defines a unique
RL signal transduction protein."
RN Nature 352:73-77(1991).
RP PHOSPHORYLATION SITES.
RX MEDLINE-94067102; PubMed=7504175;
```

RA Sun X.J., Crimmins D.L., Myers M.G., Miralpeix M., White M.F.;
RT "Pleiotropic insulin signals are engaged by multisite phosphorylation
of IRS-1";
RL Mol. Cell. Biol. 13:7418-7428(1993).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE-93352637; PubMed-8349691;
RA Tanasijevic M.J., Myers M.G., Thoma R.S., Crimmins D.L., White M.F.,
RT "Phosphorylation of the insulin receptor substrate IRS-1 by casein
kinase II";
RL J. Biol. Chem. 268:18157-18166(1993).
CC -|- FUNCTION: MAY MEDIATE THE CONTROL OF VARIOUS CELLULAR PROCESSES
BY INSULIN. WHEN PHOSPHORYLATED BY THE INSULIN RECEPTOR BINDS
SPECIFICALLY TO VARIOUS CELLULAR PROTEINS CONTAINING SH2 DOMAINS
SUCH AS PHOSPHATIDYLINOSITOL 3-KINASE P85 SUBUNIT OR GRB-2.
CC -|- SIMILARITY: CONTAINS 1 PH DOMAIN
CC -|- SIMILARITY: CONTAINS 1 PTB DOMAIN
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X58375; CAA41264.1; -
DR PIR; S16948; S16948.
DR HSSP; P35568; 1IRS.
DR InterPro; IPR002404; Insulin_Recep_S-1.
DR InterPro; IPR001849; PH.
DR Pfam; PF02174; PH; 1.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00628; INSULINRS1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTBI; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Phosphorylation.
FT DOMAIN 12 115
FT DOMAIN 152 262
FT DOMAIN 872 881
FT DOMAIN 1196 1200
FT MOD_RES 99 99
FT MOD_RES 460 460
FT MOD_RES 502 502
FT MOD_RES 608 608
FT MOD_RES 628 628
FT MOD_RES 895 895
FT MOD_RES 939 939
FT MOD_RES 987 987
FT MOD_RES 1172 1172
FT MOD_RES 1222 1222
SQ SEQUENCE 1235 AA; 131178 MW; A274BC7540CA85C5 CRC64;

Query Match 35.4%; Score 49.5; DB 1; Length 1235;
Best Local Similarity 47.8%; Pred. No. 9.9;
Matches 11; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

QY 1 CGETQSRVTH-----PHLPR 16
DB 816 CGARPESSVTHPHHALQPHLPR 838

RESULT 6
MKR3_MOUSE
ID MKR3_MOUSE STANDARD; PRT; 544 AA.
AC Q60764;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Makorin_3 (Zinc-finger protein 127).

GN MKR3 OR ZFP127 OR ZNF127.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10196368;
RA Jong M.T.C., Carey A.H., Caldwell K.A., Lau M.H., Handel M.A.,
RA Driscoll D.J., Stewart C.L., Rinchik E.M., Nicholls R.D.;
RT "Imprinting of a RING zinc-finger encoding gene in the mouse
chromosome region homologous to the Prader-Willi syndrome genetic
region.";
RL Hum. Mol. Genet. 8:795-803(1999).
CC -|- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -|- SIMILARITY: CONTAINS 3 C3H1-TYPE ZINC FINGERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U19106; AAA76863.1; -
DR MGD; MGI:99158; Zfp127.
DR InterPro; IPR000571; Zf-CCCH.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; Zf-C3HC4; 1.
DR Pfam; PF00642; Zf-CCCH; 3.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf_C3H1; 3.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
KW Zinc-finger; Repeat.
FT ZN_FING 98 116
FT ZN_FING 280 298
FT ZN_FING 302 329
FT ZN_FING 347 401
FT ZN_FING 436 456
SQ SEQUENCE 544 AA; 59444 MW; FF05B7D034C5EA9F CRC64;

Query Match 34.3%; Score 48; DB 1; Length 544;
Best Local Similarity 45.0%; Pred. No. 7.1;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPRALMRS 21
DB 236 GQYRGRVVPVPHGPEAPLQS 245

RESULT 7
GVPC_HALME
ID GVPC_HALME STANDARD; PRT; 381 AA.
AC Q02228;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gas vesicle protein C.
GN GVPC.
OS Halobacterium mediterranei (Haloflex mediterranei).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1411;
RA MEDLINE=93021102; PubMed=1404376;
RX Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
J. Mol. Biol. 227:586-592(1992).
RL


```
QY 1 CGEYQSRVTHPHL 14
   III I I : I I I I
Db 344 CGEYQQA-ITEPHL 356

RESULT 9
VMAI_INCJJ STANDARD; PRT; 242 AA.
AC P12446;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Matrix protein.
GN M.
OS Influenza C virus (strain C/JJ/50).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza C virus.
ON NCBI_TaxID=11560;
RX MEDLINE=88300888; PubMed=3404579;
RA Yanashita M., Krystal M., Palese P.;
RT "Evidence that the matrix protein of influenza C virus is coded for
   by a spliced mRNA.";
RL J. Virol. 62:3348-3355(1988).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M22038; AAA43781.1; -
DR PIR; A28878; MFIVCJ.
DR InterPro; IPR004271; CM1.
DR Pfam; PF03026; CM1; 1.
KW Matrix protein.
SQ SEQUENCE 242 AA; 26967 MW; ABF9D9054E1C9D91 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 242;
Best Local Similarity 38.9%; Pred. No. 4.3;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 CGEYQSRVTHPHLPRAL 18
   I : I : I I I :
Db 209 CNHTFGSNIMRPHLEKAI 226

RESULT 10
MYHD_HUMAN STANDARD; PRT; 1938 AA.
AC Q9UKX3; O95252;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, extraocular (MyHC-ec).
GN MYH13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Leinwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
   heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-73(1999).
```

```

RN RP SEQUENCE OF 1917-1938 FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99026150; PubMed=9806854;
RA Winters L.M., Briggs M.M., Schachar F.;
RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
   the cluster of fast and developmental myosin genes on chromosome 17.";
RL Genomics 54:188-189(1998).
-|- FUNCTION: MUSCLE CONTRACTION.
CC -|- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -|- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -|- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -|- PM: TWO CYSTEINE RESIDUES IN THE SI DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -|- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -|- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -|- SIMILARITY: CONTAINS 1 IQ DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF111782; AAD29948.1; -
DR EMBL; AF075248; AAC83241.1; -
DR HSSP; P08799; IMND.
DR MIM; 603487; -
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
FT DOMAIN 785 814 IQ.
FT DOMAIN 843 1938 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 761 775 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).
SQ SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CD5 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1938;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 RYVTHPLRALMRSTTK 24
   I I I I I I I I I I
Db 667 RSTPHFVRCLIPNETK 683
```

```

RESULT 11
MYH3_CHICK STANDARD; PRT; 1940 AA.
AC P02565;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87194881; PubMed=3571266;
RA Molina M.I., Kropp K.E., Gulick J., Robbins J.;
RT "The sequence of an embryonic myosin heavy chain gene and isolation
of its corresponding cDNA.";
RL J. Biol. Chem. 262:6478-6488(1987).
RN [2]
RP SEQUENCE OF 1502-1940 FROM N.A.
RX MEDLINE=83161144; PubMed=6833296;
RA Kavinsky C.J., Uneda P.K., Sinha A.M., Elzinga M., Tong S.W., Zak R.,
RA Jakovcic S., Rabinowitz M.;
RT "Cloned mRNA sequences for two types of embryonic myosin heavy chains
from chick skeletal muscle. I. DNA and derived amino acid sequence of
light meromyosin.";
RL J. Biol. Chem. 258:5196-5205(1983).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; J02714; AAA48972.1; -.
DR PIR; A02990; A02990.
DR PIR; A29320; A29320.
DR HSSP; P08799; LMND.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR000409; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.

```

```

DR PROSITE; PS00096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 844 1940 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 762 776 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1).
FT MOD_RES 710 710 ALKYLATION (SH-2).
FT VARIANT 379 379 G -> D.
FT CONFLICT 1547 1547 T -> A (IN REF. 2).
FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
SQ SEQUENCE 1940 AA; 222816 MW; C348333D75B04DFE2 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 52.9%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 RYTHPLPRALMRSTTK 24
| |||| | | : | |
Db 668 RSTHPLFVRCLIPNETK 684

RESULT 12
MYH3_HUMAN STANDARD; PRT; 1940 AA.
AC P11055; Q15492;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic (Muscle embryonic
DE myosin heavy chain) (SMHC).
GN MYH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263803; PubMed=2726495;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H.,
RA Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Nucleotide sequence of full length human embryonic myosin heavy
chain cDNA.";
RL Nucleic Acids Res. 17:3591-3592(1989).
RN [2]
RP SEQUENCE OF 774-1940 FROM N.A.
RX MEDLINE=90033298; PubMed=2806546;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,
RA Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Human embryonic myosin heavy chain cDNA. Interspecies sequence
conservation of the myosin rod, chromosomal locus and isoform
specific transcription of the gene.";
RL FEBS Lett. 256:21-28(1989).
RN [3]
RP SEQUENCE OF 856-1940 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90233862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [4]
RP SEQUENCE OF 856-1940 FROM N.A.
RX MEDLINE=89366648; PubMed=2771643;
RA Karsch-Mizrachi I., Travis M., Blau H., Leinwand L.A.;
RT "Expression and DNA sequence analysis of a human embryonic skeletal

```

muscle myosin heavy chain gene.";
Nucleic Acids Res. 17:6167-6179(1989).
-!- FUNCTION: MUSCLE CONTRACTION.
-!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
-!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
-!- DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE
AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL
MUSCLE.
-!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
-!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
-!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
-!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
-!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X13988; CAA32167.1; -;
DR EMBL; X13100; CAA31492.1; -;
DR EMBL; X15193; CAA35942.1; -;
DR EMBL; X15696; CAA33731.1; -;
DR PIR; S04090; S04090.
DR HSP; P08799; 1MMD.
DR MIM; 160720; -;
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
FT CONFLICT 1331 1331 A -> G (IN REF. 3).
FT CONFLICT 1391 1392 KK -> QE (IN REF. 1 AND 2).
FT CONFLICT 1508 1609 SR -> RA (IN REF. 3).
FT CONFLICT 1563 1664 RG -> QT (IN REF. 2).
SQ SEQUENCE 1940 AA; 224035 MW; 43CA58C6A4BA1253 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 47.1%; Pred. No. 39;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 8 RVTHPLPLALMRSTTK 24
| | | | | : : : |
Db 664 RTTHPEVRCIIPNETK 680

RESULT 13
MYH3_RAT
ID MYH3_RAT STANDARD; PRT; 1940 AA.
AC P12847;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, fast skeletal muscle, embryonic.
GN MYH3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060988; PubMed=3783701;
RA Strehler E.E.; Strehler-Page M.-A.; Perriard J.C.; Periasamy M.,
Nadal-Ginard B.;
RT "Complete nucleotide and encoded amino acid sequence of a mammalian
myosin heavy chain gene. Evidence against intron-dependent evolution
of the rod.";
RL J. Mol. Biol. 190:291-317(1986).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X04267; CAA27817.1; -;
DR PIR; A24922; A24922.
DR HSP; P08799; 1MMD.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
FT DOMAIN 782 811 IQ.

```
FT DOMAIN 782 811 IO.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596E5A696 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 47.1%; Pred. No. 39;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 8 RYTHPLPALMRSTTK 24
DB 664 RYTHPHFVRCIIPNETK 680

RESULT 14
IRSL_MOUSE
ID IRSL_MOUSE STANDARD; PRT; 1233 AA.
AC P35569;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Insulin receptor substrate-1.
GN IRS1 OR IRS-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=94220494; PubMed=8167159;
RA Araki E., Haag B.L. III, Kahn C.R.;
RT "Cloning of the mouse insulin receptor substrate-1 (IRS-1) gene and
RT complete sequence of mouse IRS-1."
RL Biochim. Biophys. Acta 1221:353-356(1994).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=93192326; PubMed=8448209;
RA Keller S.R., Aebersold R., Garner C.W., Lienhard G.E.;
RT "The insulin-elicited 160 kDa phosphotyrosine protein in mouse
RT adipocytes is an insulin receptor substrate 1: identification by
RT cloning."
RL Biochim. Biophys. Acta 1172:323-326(1993).
CC -!- FUNCTION: MAY MEDIATES THE CONTROL OF VARIOUS CELLULAR PROCESSES
CC BY INSULIN WHEN PHOSPHORYLATED BY THE INSULIN RECEPTOR BINDS
CC SPECIFICALLY TO VARIOUS CELLULAR PROTEINS CONTAINING SH2 DOMAINS
CC SUCH AS PHOSPHATIDYLINOSITOL 3-KINASE P85 SUBUNIT OR GRB-2.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PTB DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L24563; AAA39335.1; -.
CC DR EMBL; X69722; CAA49378.1; -.
CC DR PIR; S43514; S43514.
CC DR HSSP; P35568; 1IRS.
CC DR MGD; MGI:99454; Irs1.
CC InterPro; IPR002404; Insulin_Recep_s-1.
CC DR InterPro; IPR001849; PH.
CC DR Pfam; PF02174; IRS; 1.
CC DR Pfam; PF00169; PH; 1.
CC PRINTS; PR00628; INSULINRS1.
```

```
DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTBI; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW Phosphorylation.
FT DOMAIN 12 115 PH.
FT DOMAIN 152 262 PTB.
FT DOMAIN 675 680 POLY-SER.
FT DOMAIN 872 877 POLY-GLN.
FT DOMAIN 1119 1128 POLY-GLY.
FT DOMAIN 1194 1198 POLY-PRO.
FT MOD_RES 99 99 PHOSPHORYLATION (BY CK2)
FT MOD_RES 99 99 (BY SIMILARITY).
FT MOD_RES 460 460 PHOSPHORYLATION (BY INSR)
FT MOD_RES 608 608 (BY SIMILARITY).
FT MOD_RES 608 608 PHOSPHORYLATION (BY INSR)
FT MOD_RES 628 628 (BY SIMILARITY).
FT MOD_RES 628 628 PHOSPHORYLATION (BY INSR)
FT MOD_RES 891 891 (BY SIMILARITY).
FT MOD_RES 891 891 PHOSPHORYLATION (BY INSR)
FT MOD_RES 935 935 (BY SIMILARITY).
FT MOD_RES 935 935 PHOSPHORYLATION (BY INSR)
FT MOD_RES 983 983 (BY SIMILARITY).
FT MOD_RES 983 983 PHOSPHORYLATION (BY INSR)
FT MOD_RES 1173 1173 (BY SIMILARITY).
FT MOD_RES 1173 1173 PHOSPHORYLATION (BY INSR)
FT MOD_RES 1220 1220 (BY SIMILARITY).
FT MOD_RES 1220 1220 PHOSPHORYLATION (BY INSR)
FT CONFLICT 1038 1039 MISSING (IN REF. 2).
FT CONFLICT 1182 1182 H -> R (IN REF. 2).
FT SEQUENCE 1233 AA; 130723 MW; C0E9B2D890DADD87 CRC64;

Query Match 33.2%; Score 46.5; DB 1; Length 1233;
Best Local Similarity 43.5%; Pred. No. 29;
Matches 10; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

QY 1 CGETYQSRVTH-----PHLPR 16
DB 816 CGARPESLTHPHHHVLQPHLPR 838

RESULT 15
YK14_YEAST
ID YK14_YEAST STANDARD; PRT; 116 AA.
AC P36078;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Hypothetical 13.6 kDa protein in MDHL-VMA5 intergenic region.
GN YK1084W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
CC [1]
CC SEQUENCE FROM N.A.
CC Pohl T.M., Pohl F.M.;
CC Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z28084; CAA81922.1; -.
CC DR EMBL; Z28084; CAA81922.1; -.
CC DR PIR; S37909; S37909.
CC DR SGD; S0001567; YK1084W.
CC Hypothetical protein
CC SEQUENCE 116 AA; 13641 MW; 57464E5FD9A591F8 CRC64;
```

Query Match 32.9%; Score 46; DB 1; Length 116;
 Best Local Similarity 45.0%; Pred. No. 2.8;
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 6 QSRVTHPHLPRLMRSTTKC 25
 ||| | ||| : :
 Db 14 QSRCVHHWLPKDVIAIRFKC 33

Search completed: August 26, 2002, 10:17:02
 Job time: 698 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: August 26, 2002, 10:04:29 ; Search time 66.82 Seconds
(without alignments)
64.724 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_19:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mbc:*
 - 8: sp_organelle:*
 - 9: sp_phage:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_virus:*
 - 13: sp_vertebrate:*
 - 14: sp_unclassified:*
 - 15: sp_rvirus:*
 - 16: sp_bacteriap:*
 - 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	54	38.6	684	13 Q90544	Q90544 ginglymosto
2	53.5	38.2	333	10 Q92761	Q92761 sambucus ni
3	53.5	38.2	333	10 Q9SY54	Q9SY54 sambucus ni
4	53.5	38.2	333	10 Q9SY55	Q9SY55 sambucus ni
5	52	37.1	320	10 Q9L087	Q9L087 arabidopsis
6	52	37.1	533	2 Q9RJ70	Q9RJ70 streptomyce
7	50.5	36.1	333	10 Q92760	Q92760 sambucus ni
8	50	35.7	1103	12 Q91PA7	Q91PA7 ribgrass mo
9	50	35.7	1107	12 Q88603	Q88603 tobacco mos
10	50	35.7	1597	12 Q91PA8	Q91PA8 ribgrass mo
11	50	35.7	1601	12 Q88604	Q88604 tobacco mos
12	49.5	35.4	330	10 Q944B9	Q944B9 sambucus ni
13	49	35.0	273	5 Q9VFL6	Q9VFL6 drosophila
14	49	35.0	277	10 Q9LTK6	Q9LTK6 arabidopsis
15	49	35.0	360	5 Q9V9X5	Q9V9X5 drosophila
16	49	35.0	840	13 Q9I8N3	Q9I8N3 rana pipien

17	48	34.3	1065	10 Q9FHK6	Q9FHK6 arabidopsis
18	48	34.3	1065	10 Q489A8	Q489A8 arabidopsis
19	48	34.3	2254	10 Q9LNO2	Q9LNO2 arabidopsis
20	47	33.6	242	12 Q89465	Q89465 influenza c
21	47	33.6	242	12 Q67390	Q67390 influenza c
22	47	33.6	242	12 Q9WAD3	Q9WAD3 influenza c
23	47	33.6	242	12 Q9WAE1	Q9WAE1 influenza c
24	47	33.6	242	12 Q9IQ57	Q9IQ57 influenza c
25	47	33.6	242	12 Q9IQ55	Q9IQ55 influenza c
26	47	33.6	242	12 Q39842	Q39842 influenza c
27	47	33.6	314	11 Q9ESH2	Q9ESH2 rattus norv
28	47	33.6	374	12 Q9W841	Q9W841 influenza c
29	47	33.6	374	12 Q9W8X9	Q9W8X9 influenza c
30	47	33.6	374	12 Q9YMB6	Q9YMB6 influenza c
31	47	33.6	374	12 Q9WAD5	Q9WAD5 influenza c
32	47	33.6	374	12 Q9WAD7	Q9WAD7 influenza c
33	47	33.6	374	12 Q9WAD9	Q9WAD9 influenza c
34	47	33.6	374	12 Q9YMB5	Q9YMB5 influenza c
35	47	33.6	374	12 Q9WAE2	Q9WAE2 influenza c
36	47	33.6	374	12 Q9WAE4	Q9WAE4 influenza c
37	47	33.6	374	12 Q67388	Q67388 influenza c
38	47	33.6	374	12 Q67389	Q67389 influenza c
39	47	33.6	374	12 Q67391	Q67391 influenza c
40	47	33.6	374	12 Q67392	Q67392 influenza c
41	47	33.6	374	12 Q96792	Q96792 influenza c
42	47	33.6	840	13 Q9I8N4	Q9I8N4 rana pipien
43	47	33.6	841	13 Q9I8N5	Q9I8N5 rana pipien
44	47	33.6	1936	13 Q90YF6	Q90YF6 paracirrhith
45	47	33.6	1938	6 Q9GJP9	Q9GJP9 oryctolagus

ALIGNMENTS

RESULT 1					
Q90544					
ID Q90544	PRELIMINARY;	PRT;	684 AA.		
AC Q90544;					
DT 01-NOV-1996 (TREMBLrel. 01, Created)					
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)					
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.					
OS Ginglymostoma cirratum (Nurse shark).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;					
OC Elasmobranchii; Galeomorphii; Galeoidea; Orectolobiformes;					
OC Ginglymostomatidae; Ginglymostoma.					
OX NCBI_TaxID=7801;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=SPLEEN;					
RX MEDLINE=95183140; PubMed=7877689;					
RA Greenberg A.S.; Avila D., Hughes M., Hughes A., McKinney E.C.,					
RA Flajnik M.F.;					
RT "A new antigen receptor gene family that undergoes rearrangement and					
RT extensive somatic diversification in sharks.";					
RL Nature 374:168-173(1995).					
DR EMBL; UI8701; AAB48195.1; -.					
DR HSSP; P01709; 2MCG.					
DR InterPro; IPR003597; Ig_c1.					
DR InterPro; IPR003600; Ig_like.					
DR InterPro; IPR003006; Ig_MHC.					
DR Pfam; PF00047; ig; 6.					
DR SMART; SM00407; Igcl; 4.					
DR SMART; SM00410; IG_like; 2.					
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.					
KW Signal; Receptor.					
FT SIGNAL 1 18 . POTENTIAL.					
FT CHAIN 19 684 NOVEL ANTIGEN RECEPTOR.					
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D207ICDA6DFD CRC64;					
Query Match	38.6%;	Score 54;	DB 13;	Length 684;	
Best Local Similarity	57.1%;	Pred. No. 3.1;			

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eucrosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RC Liu S., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,
RA Toriumi M., Chin C., Chitau J., Choi E., Chung M., Gonzalez A.,
RA Huang B., Liu A., Vaysberg M., Altafi H., Brooks S., Buehler E.,
RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C., Shinn P.,
RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;
RT *The sequence of BAC T1n6 from Arabidopsis thaliana chromosome 1 *;
RT Submitted (JUN-2000) to the EMBL/GenBank/thaliana databases.
DR EMBL; ACQ09273; AAF78401.1; -;
DR HSSP; P33911; IC80.
DR InterPro; IPR003703; Acyl_COA_thio.
DR Pfam; PF02551; Acyl_COA_thio: 1.
SQ SEQUENCE 320 AA; 36157 MW; 03A82A96333E1BFC CRC64;

OC Asteridae; euasterids II; Adoxaceae; *Sambucus*.
 NCBI_TaxId=4202;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=ELDERBERRY FRUIT;
 RC MEDLINE=99214486; PubMed=10198114;
 RX
 RA Van Damme E.J.M., Charels D., Roy S., Tierens K., Barre A.,
 RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.:
 RT "A gene encoding a hevein-like protein from elderberry fruits is
 RT homologous to PR-4 and class V chitinase genes";
 RL Plant Physiol. 119:1547-1556(1999).

homologous to PR-4 and class V chitinase genes.;
 RL Plant Physiol. 119:1547-1556(1999).
 DR EMBL: AF074386; AAD11407.1; -.
 DR HSSP: P02877; 1HEV.
 DR InterPro: IPR001002; Chitin_bind.
 DR InterPro: IPR000726; Glyco_hydro_19.
 DR Pfam: PF00187; chitin_binding_1.
 DR Pfam: PF00182; Glyco_hydro_19; 1.
 DR ProDom: PD000574; Glyco_hydro_19; 1.
 DR ProDom: PD000509; Chitin_bind; 1.
 DR SMART: SM00270; ChtBD1; 1.
 DR PROSITE: PS00774; CHITINASE_19_2; 1.
 DR PROSITE: PS00026; CHITIN_BINDING; 1.
 DR Chitin-binding.
 KW
 SQ SEQUENCE 333 AA: 36922 MW: 70C6969392A1174A CRC64;

```

Query Match          36.1%; Score 50.5; DB 10; Length 333;
Best Local Similarity 40.0%; Pred. No. 5.3;
Matches 14; Conservative 3; Mismatches 3; Indels 15; Gaps:
                                0
Qy      1 CGETQQ-----SRVTHPLPLALMRST 22
      || ||| :|| |||||:| |
Db      52 CGSTQYCEDGCGSQCRDTARLT--DLPLALLRPT 84

RESULT      8
ID Q91PA7 PRELIMINARY; PRT: 1103 AA.
AC Q91PA7;
DT 01-DEC-2001 (*REMBLrel. 19, Created)
DT 01-DEC-2001 (*REMBLrel. 19, Last sequence update)

```



```

Qy      1 CGETYQ-----SRVTHPHLPALMRST 22
         || |||      |:| ||||:| |
Db      49 CGSTIYQCEDGCQSQCRDTGRLT--DLPRALLRPT 81

```

RESULT	14
Q9LTK6	
ID	Q9LTK6

```

AC Q9LRK6;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE GENOMIC DNA, CHROMOSOME 5, BAC CLONE:F17P19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB025603; BAA97457.1; -.
SQ SEQUENCE 277 AA; 31592 MW; C72F0D844D6A2089 CRC64;

Query Match 35.0%; Score 49; DB 10; Length 277;
Best Local Similarity 36.4%; Pred. No. 7.6;
Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPLPRALMRSTT 23
||| |||||
Db 148 GENFSGRYTHTHSPSSISHNPT 169

RESULT 15
Q9V9X5
ID Q9V9X5 PRELIMINARY; PRT; 360 AA.
AC Q9V9X5;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CG11335 PROTEIN (LYSYL OXIDASE-LIKE).
GN LOX OR CG11335.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```

```

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RA Mink M., Csizsar K.;
RT "The role of Drosophila lysyl oxidase in the regulation of chromatin
RT structure.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AE003778; AAF57153.1; -.
DR EMBL; AJ295625; CAB99480.1; -.
DR FlyBase; FBgn0039848; lox;
DR InterPro; IPR001695; Lysyl_oxidase.
DR InterPro; IPR001190; SRCR.
DR Pfam; PF01186; Lysyl_oxidase; 1.
DR Pfam; PF00530; SRCR; 1.
DR PRINTS; PR00074; LYSYLOXIDASE.
DR PRINTS; PR00258; SPERACTRCPTR.
DR ProDom; PD013887; Lysyl_oxidase; 1.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS0287; SRCR_2; 1.
SQ SEQUENCE 360 AA; 40832 MW; 7130F7265CB5B698 CRC64;

```

Query Match 35.0%; Score 49; DB 5; Length 360;

Best Local Similarity 40.9%; Pred. No. 10;

Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 ETYQSRVTHPLPRALMRSTTK 24

||| ||||| |||||

Db 184 DAYEIRRTNPAAHILLRFSVK 205

Search completed: August 26, 2002, 10:16:17

Job time: 708 sec

THIS PAGE BLANK (USPTO)

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 69.6%; Score 94; DB 4; Length 25;
Best Local Similarity 64.0%; Pred. No. 8e-09;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 1 CGETYQSRVTHPLRALMRSTTKC 25

RESULT 3
US-08-232-539D-19
; Sequence 19, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-19

Query Match 43.0%; Score 58; DB 2; Length 22;
Best Local Similarity 55.6%; Pred. No. 0.005;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREV 19
DB 5 GETYQCRVTHPLRALM 22

RESULT 4
US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-20

Query Match 43.0%; Score 58; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 0.0055;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREV 19
||||| :||| ||| :
Db 7 GETYQCRVTHPHLPRLM 24

RESULT 5
US-09-187-789-10
; Sequence 10, Application US/09187789
; Patent No. 6340740
; GENERAL INFORMATION:
; APPLICANT: Alnemri, Emad S.
; APPLICANT: Fernandez-Alnemri, Teresa
; TITLE OF INVENTION: CASPASE-14, AN APOTOTIC PROTEASE, NUCLEIC ACID ENCODING
; FILE REFERENCE: 480140.434C1
; CURRENT APPLICATION NUMBER: US/09/187,789
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-187-789-10

Query Match 27.4%; Score 37; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 13;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 11 HPDLPREVVRISAK 24
||| | :| :|
Db 6 HPDKPLKVLQGLK 19

RESULT 6
US-09-017-205-5
; Sequence 5, Application US/09017205
; Patent No. 5965357
; GENERAL INFORMATION:
; APPLICANT: Marsden, Howard S
; TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN
; TITLE OF INVENTION: DIAGNOSIS OF HERPES SIMPLEX VIRUS TYPE 2
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 5965357th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,205
; FILING DATE: 02-FEB-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 604-436
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide from HSV-2 glycoprotein G
; FRAGMENT TYPE: internal
US-09-017-205-5

Query Match 25.9%; Score 35; DB 2; Length 18;
Best Local Similarity 87.5%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 DLPREVVR 20
||||| |||
Db 10 DLPRRVVR 17

RESULT 7
US-09-077-991-5
; Sequence 5, Application US/09077991
; Patent No. 6207375
; GENERAL INFORMATION:
; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T.C.
; APPLICANT: Roche, P.C.
; TITLE OF INVENTION: TGF-Beta inducible early factor-1
; TITLE OF INVENTION: (TIEF-1) and a method to detect breast cancer
; FILE REFERENCE: 150.157US2
; CURRENT APPLICATION NUMBER: US/09/077,991
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: PCT/US96/19555
; EARLIER FILING DATE: 1996-12-11
; EARLIER APPLICATION NUMBER: US 08/570,227
; EARLIER FILING DATE: 1995-12-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-991-5

Query Match 25.9%; Score 35; DB 4; Length 21;
Best Local Similarity 57.1%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYKSTVSHPDLP 16
| ||| |||
Db 3 EEKSPVSAKLPK 16

RESULT 8
US-09-177-249-298
; Sequence 298, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Ohad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadegari, Ramin
; APPLICANT: Margossian, Linda
; APPLICANT: Harada, John
; APPLICANT: Goldberg, Robert B.
; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
; FILE OF INVENTION: Development in Plants
; FILE REFERENCE: 023070-086120US
; CURRENT APPLICATION NUMBER: US/09/177,249
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US 09/071,838
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 298
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-177-249-298

Query Match 23.7%; Score 32; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 66;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 16 REVVRSIAKC 25
DB 2 KSVVRNVQKC 11

RESULT 9
US-08-269-441A-15
; Sequence 15, Application US/08269441A
; Patent No. 5552529
; GENERAL INFORMATION:
; APPLICANT: Rearden, Ann
; TITLE OF INVENTION: A NOVEL AUTOANTIGEN, PINCH
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/269,441A
; FILING DATE: 30-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/009001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..11
US-08-269-441A-15

Query Match 23.0%; Score 31; DB 1; Length 11;
Best Local Similarity 45.5%; Pred. No. 53;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 YKSTVSHPDLP 15
DB 1 FKPPKYHPDVP 11
RESULT 10
US-09-078-173A-23
; Sequence 23, Application US/09078173A
; Patent No. 6200794
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskins
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; FILE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0001
; CURRENT APPLICATION NUMBER: US/09/078,173A
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Psidium Guajava (guava)
US-09-078-173A-23

Query Match 23.0%; Score 31; DB 4; Length 13;
Best Local Similarity 46.2%; Pred. No. 65;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 TYKSTVSHPDLP 16
DB 1 TYPPSLSPSPSPR 13

RESULT 11
US-09-078-173A-24
; Sequence 24, Application US/09078173A
; Patent No. 6200794
; GENERAL INFORMATION:
; APPLICANT: Ian M. Whitehead
; APPLICANT: Alan Slusarenko
; APPLICANT: Duncan Gaskins
; APPLICANT: Alan Brash
; APPLICANT: Nathalie Tijet
; TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
; FILE OF INVENTION: LYASE AND USES THEREOF
; FILE REFERENCE: 06027.0001
; CURRENT APPLICATION NUMBER: US/09/078,173A
; CURRENT FILING DATE: 1998-05-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Psidium Guajava (guava)
US-09-078-173A-24

Query Match 23.0%; Score 31; DB 4; Length 14;
Best Local Similarity 46.2%; Pred. No. 71;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 4 TYKSTVSHPDLP 16
DB 1 TYPPSLSPSPSPR 13

RESULT 12
US-07-976-358-24

```
; Sequence 24, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-976-358-24

Query Match 23.0%; Score 31; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 YKSTVSHPDLPREV 18
| | : | | | |
Db 7 YSSQTLPLSPRVV 20

RESULT 13
US-07-976-358-15
; Sequence 15, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
```

```
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-976-358-15

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 YKSTVSHPDLPREV 18
| | : | | | |
Db 8 YSSQTLPLSPRVV 21

RESULT 14
US-07-976-358-18
; Sequence 18, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-976-358-18

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 52.9%; Pred. No. 1.4e+02;
Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

QY 5 YKSTVSHPOLPREVRS 21
| | | | | | | | |
Db 8 YSSHPLPLPRAVPS 24

RESULT 15
US-07-976-358-21
; Sequence 21, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHODYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-976-358-21

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 YKSTVSHPOLPREV 18
| | | | | | | | |
Db 8 YSSQSTLPSLPRV 21

Search completed: August 26, 2002, 10:19:37
Job time: 393 sec

us-09-701-623c-8.closed.rpr

Mon Aug 26 11:09:38 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model
Run on: August 26, 2002, 10:20:42 ; Search time 43.94 seconds
(without alignments)
54.671 Million cell updates/sec

Title: US-09-701-623c-8
Perfect score: 140
Sequence: 1 CGYGYSIVDRDPFKPIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 4981

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	26.4	23	S65374	cytochrome-c oxida
2	33	23.6	18	B49254	TCR C gamma 1 chai
3	30	21.4	15	PS0450	23K protein 4307
4	30	21.4	20	S44465	pyrrhocoricin - Py
5	30	21.4	22	S73389	hypothetical prote
6	29	20.7	17	D48138	d(TTAGGG)n-binding
7	28.5	20.4	23	E39855	paralytic peptide
8	28.5	20.4	23	D39855	paralytic peptide
9	28	20.0	13	S21152	tryptophyllin-rela
10	28	20.0	17	S33609	extensin - maize (
11	28	20.0	22	S52357	hypothetical prote
12	28	20.0	23	A32473	histidine-rich pro
13	28	20.0	23	JP0045	ribosomal protein
14	27	19.3	21	S47207	T-cell receptor J-
15	27	19.3	25	B57001	endo-1,4-beta-xyla
16	26.5	18.9	24	T42257	phosphoprotein pho
17	26	18.6	11	PN0042	stathmin - mouse (
18	26	18.6	13	A05174	tryptophyllin-13 -
19	26	18.6	15	PA0088	protein QF200051 -
20	26	18.6	22	PN0910	N4-(beta-N-acetyl)g
21	26	18.6	23	C39855	paralytic peptide
22	26	18.6	23	B61079	listeriolysin O -
23	26	18.6	25	I40692	cenA protein (Iga1
24	25	17.9	14	S50900	chlorophyll a/b-bi
25	25	17.9	19	S59485	hydroxyproline-rich
26	25	17.9	22	S55236	T669 kinase - huma
27	25	17.9	22	S37452	gene E6 protein (c
28	25	17.9	23	A56357	tyrosine kinase su
29	25	17.9	24	S53793	actin - mouse (fra

ALIGNMENTS

RESULT 1

S65374 cytochrome-c oxidase (EC 1.9.3.1) chain IV, cardiac - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
C:Accession: S65374

R:Schaeffer, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.
Eur. J. Biochem. 230, 235-241, 1995

A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-t

A:Reference number: S65372; MUID:95324529

A:Accession: S65374

A:status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <SCH>

C:Superfamily: cytochrome-c oxidase chain IV

C:Keywords: cardiac muscle; heart; oxidoreductase

Query Match 26.4%; Score 37; DB 2; Length 23;
Best Local Similarity 53.8%; Pred. No. 38;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 YGYOSIVDRDPDP 15
DB 11 YALPSYVDRDYP 23

RESULT 2

B49254 TCR C gamma 1 chain V-J region - mouse (fragment)

C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Jul-1999

C:Accession: B49254

R:Ezquerria, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E

Eur. J. Immunol. 22, 491-498, 1992

A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of th

A:Reference number: A49037; MUID:92164730

A:Accession: B49254

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-18 <EQ>

A:Cross-references: GB:S90688; NID:g246308; PIDN:AAB21557.1; PID:g246309

A:Experimental source: Y93A cells

A>Note: sequence extracted from NCBI backbone (NCBIN:90688, NCBIP:90690)

C:Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 23.6%; Score 33; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYGYQS 7

Db 5 CSYGYS 11

RESULT 3

23K protein 4307 - rice (strain Nihonbare) (fragment)
 C:Species: Oryza sativa (rice)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
 R:Tsugita, A.; Miyatake, N.
 submitted to JIPID, April 1993
 A:Accession: PS0450
 A:Reference number: PS0208
 A:Molecule type: protein
 A:Residues: 1-15 <TSU>
 A:Experimental source: germ, strain Nihonbare
 C:Comment: molecular weight 23K, pI 5.3.

Query Match 21.4% Score 30; DB 2; Length 15;
 Best Local Similarity 66.7% Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 YQSIYDRPD 13
 ||| ||| |
 Db 2 YQIDPRKD 10

RESULT 4

S44465
 Pyrithocorin - Pyrithocoris apterus
 C:Species: Pyrithocoris apterus
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
 R:Cociancich, S.; Dupont, A.; Hegy, G.; Lanot, R.; Holder, F.; Hetru, C.; Hoffmann, J.A.
 Biochem. J. 300: 567-575, 1994
 A:Title: Novel inducible antibacterial peptides from a hemipteran insect, the sap-sucking
 A:Reference number: S44465; MUID:94271176
 A:Accession: S44465
 A:Molecule type: protein
 A:Residues: 1-20 <COC>
 C:Function:
 A:Description: antibacterial protein
 A:Note: active against gram-negative bacteria
 C:Keywords: antibacterial; hemolymph; immune response

Query Match 21.4% Score 30; DB 2; Length 20;
 Best Local Similarity 50.0% Pred. No. 3.9e+02;
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 SIVDRPDPRKI 18
 | | | | | | |
 Db 5 SYLPRPTPRPI 16

RESULT 5

S73389
 hypothetical protein A05_Orf139 - Mycoplasma pneumoniae (strain ATCC 29342)
 C:Species: Mycoplasma pneumoniae
 A:Variety: ATCC 29342
 C:Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
 R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24: 4420-4449, 1996
 A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
 A:Reference number: S73389; MUID:97105885
 A:Accession: S73389
 A>Status: preliminary
 A:Molecule type: nucleic acid sequence not shown; translation not shown
 A:Residues: 1-22 <HIM>
 A:Cross-references: EMBL:AE000008; GB:U00089; NID:g1673711; PIDN:AAB95715.1; PID:g1673711

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
 C:Genetic code: SGC3
 A:Start codon: GTG

Query Match 21.4% Score 30; DB 2; Length 22;
 Best Local Similarity 50.0% Pred. No. 4.4e+02;
 Matches 8; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
 QY 2 GYGYSIVDR--PDFF 15
 | | | | | | | |
 Db 6 GSPYHWNLDLRNPDFF 21

RESULT 6

D48138
 d(TTAGGG)n-binding protein B37 - human (fragment)
 N:Alternate names: type A-B heterogeneous nuclear ribonucleoprotein homolog
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997
 R:Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.
 Mol. Cell. Biol. 13, 4301-4310, 1993
 A:Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) an
 A:Reference number: A48138; MUID:93309464
 A:Accession: D48138
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-17 <ISH>
 A:Experimental source: HeLa cell nuclei
 A:Note: sequence extracted from NCBI backbone (NCBIP:134642)

Query Match 20.7% Score 29; DB 2; Length 17;
 Best Local Similarity 38.5% Pred. No. 4.7e+02;
 Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 YQSIYDRPDPRKP 17
 | | | | | | |
 Db 1 FGEVVDXTDTPD 13

RESULT 7

E39855
 paralytic peptide III - beet armyworm
 C:Species: Spodoptera exigua (beet armyworm)
 C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
 R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.
 J. Biol. Chem. 266, 12873-12877, 1991
 A:Title: Isolation and identification of paralytic peptides from hemolymph of the lep
 A:Reference number: A39855; MUID:91302298
 A:Accession: E39855
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-23 <SKI>
 C:Superfamily: paralytic peptide I

Query Match 20.4% Score 28.5; DB 2; Length 23;
 Best Local Similarity 41.2% Pred. No. 7.8e+02;
 Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYSIVD---RPDF 14
 | | | | | | |
 Db 7 CTPGYORTADGRCKPTF 23

RESULT 8

D39855
 paralytic peptide II - beet armyworm
 C:Species: Spodoptera exigua (beet armyworm)

C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: D39855
R;Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.
J. Biol. Chem. 266, 12873-12877, 1991
A;Title: Isolation and identification of paralytic peptides from hemolymph of the lepidopteran *Phthorimaea operculella* (Lepidoptera: Tortricidae)
A;Reference number: A39855; MUID:91302298
A;Accession: D39855
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-23 <SKI>
C;Superfamily: paralytic peptide I

Query Match 20.4%; Score 28.5; DB 2; Length 23;
Best Local Similarity 41.2%; Pred. No. 7.8e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYQSIVD---RPDF 14
| | | | | : | |
Db 7 CTPGYQRTADGRCKPTF 23

RESULT 9
S21152
tryptophyllin-related peptide - two-colored leaf frog
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 18-Aug-2000
C;Accession: S21152
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of the frog *Phyllomedusa bicolor* (Anura: Phyllomedusidae)
A;Reference number: S21152; MUID:92339502
A;Accession: S21152
A;Molecule type: protein
A;Residues: 1-13 <MIG>
A;Experimental source: skin
C;Superfamily: unassigned animal peptides

Query Match 20.0%; Score 28; DB 2; Length 13;
Best Local Similarity 44.4%; Pred. No. 4.9e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 DRDPFKPI 18
: | : | | |
Db 2 EKFPYPPPI 10

RESULT 10
S33609
extensin - maize (fragment)
C;Species: Zea mays (maize)
C;Date: 19-Mar-1997 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C;Accession: S33609
R;Murphy, J.M.; Hood, E.E.
Plant Mol. Biol. 21, 885-893, 1993
A;Title: Molecular basis for extensin size heterogeneity in two maize varieties.
A;Reference number: S33609; MUID:93222485
A;Accession: S33609
A;Molecule type: protein
A;Residues: 1-17 <MUR>
C;Keywords: glycoprotein; hydroxyproline

Query Match 20.0%; Score 28; DB 2; Length 17;
Best Local Similarity 42.1%; Pred. No. 6.7e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 8; Gaps 2;

QY 2 GYGYQSIVDRPDF---PKP 17
| | | | | : | | | |
Db 3 GYGY-----GPPYTPPKP 16

RESULT 11
S52357
hypothetical protein - human
C;Species: Homo sapiens (man)
C;Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C;Accession: S52357
R;Agryrokastritis, A.; Leversha, M.A.; Ferguson-Smith, M.; Moschonas, M.K.
submitted to the EMBL Data Library, March 1993
A;Description: A cosmid clone mapped to human chromosome 11p15 detects a Taq I restrl
A;Reference number: S52355
A;Accession: S52357
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-22 <AGR>
A;Cross-references: EMBL:X72881; NID:g667002; PIDN:CAA51393.1; PID:g667004

Query Match 20.0%; Score 28; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 8.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 VDREDF 14
: | | | | :
Db 1 MDRPDY 6

RESULT 12
A32473
histidine-rich protein C - liver fluke (fragment)
C;Species: Fasciola hepatica (liver fluke)
C;Date: 25-Sep-1989 #sequence_revision 12-May-1994 #text_change 15-Oct-1999
C;Accession: A32473
R;Waite, J.H.; Rice-Ficht, A.C.
Biochemistry 28, 6104-6110, 1989
A;Title: A histidine-rich protein from the vitellaria of the liver fluke *Fasciola hepatica*
A;Reference number: A32473; MUID:89375343
A;Accession: A32473
A;Molecule type: protein
A;Residues: 1-23 <WAI>
A;Note: 22-Gly, 22-Ser, 23-Gly, and 23-Ser were also found
C;Keywords: egg yolk
F;6,8,12,16/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental

Query Match 20.0%; Score 28; DB 2; Length 23;
Best Local Similarity 83.3%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGYQS 7
| | | | |
Db 5 GYGYGS 10

RESULT 13
JP0045
ribosomal protein L30 - Staphylococcus aureus (fragment)
C;Species: Staphylococcus aureus
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 03-May-1996
C;Accession: JP0045
R;Ochi, K.
submitted to JIPID, February 1994
A;Description: Phylogenetic diversity in the genus *Bacillus* and comparative ribosomal
A;Reference number: JP0042
A;Accession: JP0045
A;Molecule type: protein
A;Residues: 1-25 <OCH>
C;Superfamily: Escherichia coli ribosomal protein L30
C;Keywords: protein biosynthesis; ribosome

Query Match 20.0%; Score 28; DB 2; Length 25;
Best Local Similarity 23.5%; Pred. No. 1e+03;
Matches 4; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

```

Qy      6 QSIVDRDPFKPIVRSI 22
       :|::||:::|:::
Db      9 RSVIGRPETQRKTEAL 25

RESULT 14
S47207
T-cell receptor J-alpha wvVIII.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: S47207
R:Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
Submitted to the EMBL Data Library, February 1993
A:Reference number: S40153
A:Accession: S47207
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-21 <PLA>
A:Cross-references: EMBL:X71032; NID:g506610; PID:g510318
C:Keywords: T-cell receptor

```

```
Query Match          19.3%; Score 27; DB 2; Length 21;
Best Local Similarity 55.6%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

Qy 7 SIVDRPDFP 15
|:| | |
Db 10 SLVPEPDCP 18

RESULT 15
B57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 2 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 06-Dec-1996
C:Accession: B57001
R:Grabski, A. C.; Forrester, I. T.; Patel, R.; Jeffries, T. W.
Protein Expr. Purif. 4, 120-129, 1993
A:title: Characterization and N-terminal amino acid sequences of beta-(1-4) endoxylanases
A:Reference number: A57001; MUID:93229899
A:Accession: B57001

Query Match 19.3%; Score 27; DB 2; Length 25;
Best Local Similarity 41.7%;
Pred. No. 1.5e+03;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GYGYQSIVDRPD 13
 ||| | | |
 Db 14 GYYSEFXTDAPN 25

Search completed: August 26, 2002, 10:20:42
Job time: 348 sec

RA Chernysh S., Cociancich S., Briand J.-P., Het

```
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
CC Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

Query Match 23.9%; Score 33.5; DB 1; Length 15;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 9 VDRPDF-PKP 17
DB 1 VDKPDYRPRP 10

RESULT 3
ID MK2A_PALPR STANDARD; PRT; 15 AA.
AC P80409;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Metalnikowin IIA.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomioidea;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE-Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
CC Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1893 MW; 23835D063B946299 CRC64;

Query Match 23.9%; Score 33.5; DB 1; Length 15;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 9 VDRPDF-PKP 17
DB 1 VDKPDYRPRP 10

RESULT 4
ID MK3_PALPR STANDARD; PRT; 16 AA.
AC P80411;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Metalnikowin III.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomioidea;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
```

```
RP SEQUENCE.
RC TISSUE-Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
CC Antibiotic; Insect immunity.
SQ SEQUENCE 16 AA; 2024 MW; A9E3835D063B9462 CRC64;

Query Match 23.9%; Score 33.5; DB 1; Length 16;
Best Local Similarity 60.0%; Pred. No. 43;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 9 VDRPDF-PKP 17
DB 1 VDKPDYRPRP 10

RESULT 5
ID PYRR_PYRAP STANDARD; PRT; 20 AA.
AC P37362; P80307;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pyrrhocoricin.
OS Pyrrhocoris apterus (Sap sucking bug).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Lygaeoidea;
OC Pyrrhocoridae; Pyrrhocoris.
OX NCBI_TaxID=37000;
RN [1]
RP SEQUENCE.
RC TISSUE-Hemolymph;
RA MEDLINE=94271176; PubMed=8002963;
RA Cociancich S., Dupont A., Hegy G., Lanot R., Holder F., Hetru C.,
RA Hofmann J.A., Bulet P.;
RT "Novel inducible antibacterial peptides from a hemipteran insect, the
RT sap-sucking bug Pyrrhocoris apterus.";
RL Biochem. J. 300:567-575(1994).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE THR-11.
RX MEDLINE=99177428; PubMed=10076062;
RA Hofmann R., Bulet P., Urge L., Otvoes L. Jr.;
RT "Range of activity and metabolic stability of synthetic antibacterial
RT glycopeptides from insects.";
RL Biochim. Biophys. Acta 1426:459-467(1999).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE. AFFECTS GRAM-NEGATIVE BACTERIA
CC E. COLI 1106, P. AERUGINOSA, E. COLI D22 AND E. CLOACAE AND
CC GRAM-POSITIVE BACTERIA M. LUTEUS AND B. SUBTILIS.
CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
DR PIR; S44465; S44465.
KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
FT CARBOHYD 11 11 O-LINKED (GALNAC...);
SQ SEQUENCE 20 AA; 2341 MW; F4320EC2FF29462C CRC64;

Query Match 21.4%; Score 30; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 SIVDRPDFPKPI 18
DB 5 SYLPRPTPRPI 16
```

RESULT 6
CP23_SPOER STANDARD; PRT; 23 AA.
AC P56683;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cardioactive peptide CAP23.
OS Spodoptera eridania (Southern armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=37547;
RN [1]
RP SEQUENCE.
RX MEDLINE=99196260; PubMed=10098624;
RA Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H.,
RA Shabanowitz J., Hunt D.F., Schooley D.A.;
RT "A cardioactive peptide from the southern armyworm, Spodoptera
RT eridania".
RL Peptides 20:53-61(1999).
CC -!- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM
CC LARVAL MANDUCA SEXTA, CAUSING AN INOTROPIC EFFECT AT LOW
CC CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS
CC AT HIGH DOSES.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.
DR HSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2519 MW; 0A96D72A70855AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.6e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;
QY 1 CGVGQSIQV---RPDF 14
DB 7 CTFGQRTADGRCKPTF 23

RESULT 7
PAP2_SPOEX STANDARD; PRT; 23 AA.
AC P30256;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Paralytic peptide II (pp II).
OS Spodoptera exigua (Beet armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
RP SEQUENCE.
RX TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens".
RL J. Biol. Chem. 266:12873-12877(1991).
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.
DR PIR; E39855; E39855.
DR HSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2505 MW; 0A96CB5EB7D55AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.6e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;
QY 1 CGVGQSIQV---RPDF 14
DB 7 CTFGQRTADGRCKPTF 23

RESULT 9
TP13_PHYRO STANDARD; PRT; 13 AA.
AC P04096;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Tryptophyllin-13.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Phyllomedusa.
OX NCBI_TaxID=8394;

DR Pfam; PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2477 MW; 0A96CB4600855AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.6e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;
QY 1 CGVGQSIQV---RPDF 14
DB 7 CTFGQRTADGRCKPTF 23

RESULT 8
PAP3_SPOEX STANDARD; PRT; 23 AA.
AC P30257;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Paralytic peptide III (pp III).
OS Spodoptera exigua (Beet armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
RP SEQUENCE.
RX TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens".
RL J. Biol. Chem. 266:12873-12877(1991).
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.
DR PIR; E39855; E39855.
DR HSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2505 MW; 0A96CB5EB7D55AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.6e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;
QY 1 CGVGQSIQV---RPDF 14
DB 7 CTFGQRTADGRCKPTF 23

RESULT 9
TP13_PHYRO STANDARD; PRT; 13 AA.
AC P04096;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Tryptophyllin-13.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Phyllomedusa.
OX NCBI_TaxID=8394;

RA Montecuccchi P.C., Gozzini L., Ersperer V.;
RT "Primary structure determination of a tryptophan-containing
RT tridecapeptide from Phyllomedusa rohdei";
RL Int. J. Pept. Protein Res. 27:175-182(1986).
CC -!- TISSUE SPECIFICITY: SECRETED.
DR PIR; A05174; A05174.
KW Amphibian skin.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 13 AA; 1646 MW; 33BF33A212227773 CRC64;

Query Match 18.6%; Score 26; DB 1; Length 13;
Best Local Similarity 44.4%; Pred. No. 4.7e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 DRDPFKPI 18
DB 2 ERPYWPPPI 10

RESULT 10
PAP1_SPOEX
ID PAP1_SPOEX STANDARD; PRT; 23 AA.
AC P30355;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Paralytic peptide I (pp I).
OS Spodoptera exigua (Beet armyworm).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
CC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=71107;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=20711576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.
DR PIR; C39855; C39855.
DR HSSP; 061704; 1B5N.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2451 MW; 0A96D1F600855AE0 CRC64;

Query Match 18.6%; Score 26; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 8.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYGYQSIVD 10
DB 7 CTGYQRTAD 16

RESULT 11
CAMT_PINPS
ID CAMT_PINPS STANDARD; PRT; 24 AA.
AC P81081;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Probable caffeoyl-CoA O-methyltransferase (EC 2.1.1.104) (Trans-
DE caffeoyl-CoA 3-O-methyltransferase) (CCOAMT) (CCOAMT) (Water stress
DE responsive protein 13) (Fragments).
OS Pinus pinaster (Maritime pine).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=98418576; PubMed=9747804;
RA Costa P., Bahrman N., Frigerio J.-M., Kremer A., Plomion C.;
RT "Water-deficit-responsive proteins in maritime pine";
RL Plant Mol. Biol. 38:587-596(1998).
RN [2]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins";
RL Electrophoresis 20:1098-1108(1999).
CC -!- FUNCTION: INVOLVED IN THE REINFORCEMENT OF THE PLANT CELL WALL
CC UNDER CONDITIONS THAT TRIGGER THE DISEASE RESISTANCE RESPONSE (BY
CC SIMILARITY).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + caffeoyl-CoA = S-
CC adenosyl-L-homocysteine + feruloyl-CoA.
CC -!- PATHWAY: LIGNIN BIOSYNTHESIS.
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC IS: 5.2, ITS MW IS: 30 kDa.
KW Lignin biosynthesis; Transferase; Methyltransferase.
FT NON_TER 1 1
FT NON_TER 9 10
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2447 MW; AAB842F9EB0CF0 CRC64;

Query Match 18.6%; Score 26; DB 1; Length 24;
Best Local Similarity 44.4%; Pred. No. 9.1e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 17 PIVRSITLC 25
DB 16 PVGDGVTLIC 24

RESULT 12
PAP1_MANSE
ID PAP1_MANSE STANDARD; PRT; 23 AA.
AC P30253;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Paralytic peptide I (pp I).
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
CC Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=20711576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO

```

CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -1- SIMILARITY: BELONGS TO THE GBP / PSPI / PARALYTIC PEPTIDE FAMILY.
DR PIR; A39855; A39855.
DR HSSP; O61704; 1B5N.
DR InterPro: IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2436 MW; 0B26CB5C29855FE4 CRC64;

Query Match 17.5%; Score 24.5; DB 1; Length 23;
Best Local Similarity 35.3%; Pred. No. 1.5e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 CGYGYQIVD---RPDF 14
   | | | | |
Db 7 CATGYLRTADGRCKPTF 23

RESULT 13
FAR6_PANRE STANDARD; PRT; 10 AA.
AC P82660;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE FMRFamide-like neuropeptide PF6 (NGAPQPFVRF-amide).
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX NCBI_TaxID=6233;
RN [1]
RP SEQUENCE, FUNCTION, AND AMIDATION.
RA Moffett C.L., Marks N.J., Halton D.W., Thomson D.P., Geary T.G.,
RA Maule A.G.;
RT "Isolation, characterization and pharmacology of FMRFamide-related
RT peptides (FarPs) from free-living nematode, Panagrellus redivivus.";
RL Submitted (JUL-2000) to the SWISS-PROT data bank.
CC -1- FUNCTION: MYOACTIVE.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1132 MW; CB13E4C9D776C76D CRC64;

Query Match 17.1%; Score 24; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 7.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVR 20
   | | | |
Db 4 PQPFVR 9

RESULT 14
CALI_CALGI STANDARD; PRT; 14 AA.
AC P20728;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Calotropin DI (EC 3.4.22.-) (Fragment).
OS Calotropis gigantea (Madur) (Bowstring hemp).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Asclepiadoideae;
OC Asclepiadeae; Calotropis.
OX NCBI_TaxID=4066;
RN [1]
RP SEQUENCE.

```

```

RA Bhattacharya D., Sengupta A., Sinha N.K.;
RT "Chemical modification and amino terminal sequence of calotropin DI
RT from Calotropis gigantea.";
RL Phytochemistry 26:633-636(1987).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
CC PAPAIN FAMILY OF THIOL PROTEASES.
DR PIR; PT0026; PT0026.
DR MEROPS; C01.011; -.
DR InterPro: IPR000169; Thiolprot_act_site.
DR PROSITE; PS00139; THIOL_PROTEASE_CYS; PARTIAL.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; PARTIAL.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; PARTIAL.
KW Hydrolase; Thiol protease.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1715 MW; D993F0276CDA4662 CRC64;

Query Match 17.1%; Score 24; DB 1; Length 14;
Best Local Similarity 60.0%; Pred. No. 1e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 RPDFF 15
   | | | |
Db 2 RPEYP 6

RESULT 15
YMOR_PSEPU STANDARD; PRT; 16 AA.
AC Q02210;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in MORA 3'region (Fragment).
OS Pseudomonas putida.
OG Plasmid pMDH7.2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M10.
RX MEDLINE=93199531; PubMed=8452544;
RA Willey D.L., Caswell D.A., Lowe C.R., Bruce N.C.;
RT "Nucleotide sequence and over-expression of morphine dehydrogenase, a
RT plasmid-encoded gene from Pseudomonas putida M10.";
RL Biochem. J. 290:539-544(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M94775; AAB17357.1; -.
DR PIR; S30384; S30384.
KW Hypothetical protein; plasmid.
FT NON_TER 16 16
SQ SEQUENCE 16 AA; 1962 MW; A2F1EB8C172766ED CRC64;

Query Match 17.1%; Score 24; DB 1; Length 16;
Best Local Similarity 40.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 11 RPDFFRPDIVRSITIC 25
   | | | | |
Db 3 RQDLPRS--RSLRSC 15

```

Search completed: August 26, 2002, 10:30:18
Job time: 638 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:29:35 ; Search time 62.22 Seconds
(without alignments)
69.509 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CCGYQSIIVDRDPFKPIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 8561

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	35	25.0	21	Q9X3G4	Q9X3G4 prochloroco
2	34	24.3	25	O67969	O67969 prochloroco
3	33	23.6	21	Q9X3F2	Q9X3F2 prochloroco
4	33	23.6	21	Q9X3G0	Q9X3G0 prochloroco
5	33	23.6	23	Q95E38	Q95E38 chlamydomon
6	32	22.9	21	Q9WV25	Q9WV25 prochloroco
7	32	22.9	21	Q9WBH5	Q9WBH5 streptococc
8	32	22.9	22	Q9QV92	Q9QV92 rattus sp.
9	32	22.9	23	Q85496	Q85496 bovine leuk
10	31	22.1	14	P82340	P82340 pisum sativ
11	31	22.1	18	Q56610	Q56610 vibrio chol
12	31	22.1	18	Q95JA2	Q95JA2 sus scrofa
13	31	22.1	21	Q9X3H0	Q9X3H0 prochloroco
14	30	21.4	21	Q9X3J8	Q9X3J8 prochloroco
15	29.5	21.1	24	Q9BM10	Q9BM10 spongilla l
16	29	20.7	17	Q9UCE7	Q9UCE7 homo sapien

17	29	20.7	22	4	Q96JA7	Q96JA7 homo sapien
18	28	20.0	20	15	O41542	O41542 human immun
19	28	20.0	21	2	Q9X3E8	Q9X3E8 prochloroco
20	28	20.0	21	2	Q9X3G6	Q9X3G6 prochloroco
21	28	20.0	21	2	Q9X3K2	Q9X3K2 prochloroco
22	28	20.0	21	2	O9ZG55	O9ZG55 chlamydia t
23	28	20.0	22	4	Q13659	Q13659 homo sapien
24	28	20.0	24	12	O91KC7	O91KC7 hepatitis c
25	27	19.3	18	2	Q9X3E9	Q9X3E9 prochloroco
26	27	19.3	19	2	O9R7L3	O9R7L3 prochloroco
27	27	19.3	20	2	O67964	O67964 prochloroco
28	27	19.3	21	2	O9WV38	O9WV38 prochloroco
29	27	19.3	21	2	O9WV36	O9WV36 prochloroco
30	27	19.3	21	2	O9WV23	O9WV23 prochloroco
31	27	19.3	21	2	O9WVY2	O9WVY2 prochloroco
32	27	19.3	21	2	O9WVW5	O9WVW5 prochloroco
33	27	19.3	21	2	O9R2Y1	O9R2Y1 prochloroco
34	27	19.3	21	2	Q9X3D8	Q9X3D8 prochloroco
35	27	19.3	21	2	Q9X3H5	Q9X3H5 prochloroco
36	27	19.3	21	2	O9X3I1	O9X3I1 prochloroco
37	27	19.3	21	2	O9X3I8	O9X3I8 prochloroco
38	27	19.3	21	2	O9X3K9	O9X3K9 prochloroco
39	27	19.3	21	2	O9X3K9	O9X3K9 prochloroco
40	27	19.3	21	10	O41559	O41559 triticum ae
41	27	19.3	25	11	O41564	O41564 triticum ae
42	26.5	18.9	16	11	O63993	O63993 rattus sp.
43	26.5	18.9	22	12	O54894	O54894 mus musculu
44	26.5	18.9	24	5	O94374	O94374 bovine papl
45	26	18.6	11	5	O9NL65	O9NL65 caenorhabdi
						Q9NL65 ascaris suu

ALIGNMENTS

RESULT 1

Q9X3G4 ID Q9X3G4 PRELIMINARY; PRT; 21 AA.
AC Q9X3G4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070158; AAD23203.1; -;
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2340 MW; CC947A36F12C854F CRC64;

Query Match 25.0%; Score 35; DB 2; Length 21;
Best Local Similarity 46.7%; Pred. No. 1.5e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 7 SIVDRDPFKPIVRS 21
||: ||: ||: ||:
DB 2 SILKKPDLTDPKLR 16

RESULT 2

O67969 ID O67969 PRELIMINARY; PRT; 25 AA.
AC O67969;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIT9303;
RX MEDLINE=98123172; PubMed=9452521;
RA Urbach E., Scanlan D.J., Distel D.L., Waterbury J.B., Chisholm S.W.;
RT "Rapid diversification of marine picoplankton with dissimilar
RT light-harvesting structures inferred from sequences of Prochlorococcus
RT and Synchococcus (Cyanobacteria).";
RL J. Mol. Evol. 46:188-201(1998).
DR EMBL; AF001491; AAC05629.1; -;
FT NON_TER 25
SQ SEQUENCE 25 AA; 2833 MW; 504CCD63AC947A21 CRC64;

Query Match 24.3%; Score 34; DB 2; Length 25;
Best Local Similarity 42.9%; Pred. No. 2.6e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 8 IVDRPDPFKPIVRS 21

I:::| | | |

Db 3 ILKKPDSFDPKLR 16

RESULT 3

ID Q9X3F2 PRELIMINARY; PRT; 21 AA.

AC Q9X3F2; 21 AA.

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).

GN PETD.

OS Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI_TaxID=1220;

RN [1]

RP SEQUENCE FROM N.A.

RA Urbach E., Chisholm S.W.;

RT "Genetic diversity in Prochlorococcus populations flow cytometrically

RT sorted from the Sargasso Sea and Gulf Stream.";

RL Limnol. Oceanog. 43:1615-1630(1998).

DR EMBL; AF070149; AAD20770.1; -;

FT NON_TER 21

SQ SEQUENCE 21 AA; 2338 MW; CC92DAFF3434E54F CRC64;

Query Match 23.6%; Score 33; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 3e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21

I:::| | | |

Db 2 SVIKKPDLPKLR 16

RESULT 4

ID Q9X3G0 PRELIMINARY; PRT; 21 AA.

AC Q9X3G0; 21 AA.

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).

GN PETD.

OS Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;

RT "Genetic diversity in Prochlorococcus populations flow cytometrically

RT sorted from the Sargasso Sea and Gulf Stream.";

RL Limnol. Oceanog. 43:1615-1630(1998).

DR EMBL; AF070155; AAD20780.1; -;

FT NON_TER 21

SQ SEQUENCE 21 AA; 2324 MW; D3AFDAECFA34E54F CRC64;

Query Match 23.6%; Score 33; DB 2; Length 21;

Best Local Similarity 33.3%; Pred. No. 3e+02;

Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21

I:::| | | |

Db 2 SVIKKPDLPKLR 16

RESULT 5

ID Q95E38 PRELIMINARY; PRT; 23 AA.

AC Q95E38; 23 AA.

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).

GN PETD.

OS Chlamydomonas subcaudata.

OG Chloroplast.

OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;

OC Chlamydomonadaceae; Chlamydomonas.

OX NCBI_TaxID=163303;

RN [1]

RP SEQUENCE FROM N.A.

RA Gudyanaite-Savitch L., Kohalmi S.E., Huner N.P.A.;

RT "Structural analysis of cytochrome f from the psychrophilic alga

RT Chlamydomonas subcaudata.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY039799; AAK64210.1; -;

FT NON_TER 23

SQ SEQUENCE 23 AA; 2429 MW; F25EB6AE00DA2134 CRC64;

Query Match 23.6%; Score 33; DB 8; Length 23;

Best Local Similarity 26.7%; Pred. No. 3.3e+02;

Matches 4; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21

I:::| | | |

Db 2 SVTKKPDLPKLR 16

RESULT 6

ID Q9WVZ5 PRELIMINARY; PRT; 21 AA.

AC Q9WVZ5; 21 AA.

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).

GN PETD.

OS Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI_TaxID=1220;

RN [1]

RP SEQUENCE FROM N.A.

```

RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL; AF070172; AAD20786.1; -.
DR EMBL; AF070170; AAD20783.1; -.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2294 MW; CC947A26F034E54F CRC64;

Query Match 22.9%; Score 32; DB 2; Length 21;
Best Local Similarity 40.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 7 SIVDRDPFPRPIVRS 21
  ||: ||| | :|:
Db 2 SILKKPDLADPKLRA 16

RESULT 7
ID Q9MBH5 PRELIMINARY; PRT; 21 AA.
AC Q9MBH5; Q9MBY6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CELL WALL LYTIC ENZYME (FRAGMENT).
GN MML.
OS Streptococcus pneumoniae bacteriophage MM1.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=120574;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20392133; PubMed=10933687;
RA Gindreau E., Lopez R., Garcia P.;
RT "MM1, a temperate bacteriophage of the type 23F Spanish/USA
RT multiresistant epidemic clone of Streptococcus pneumoniae: structural
RT analysis of the site-specific integration system.";
RL J. Virol. 74:7803-7813(2000).
DR EMBL; AJ400632; CAB96625.1; -.
DR EMBL; AJ400629; CAB96617.1; -.
FT NON_TER 1 1
SQ SEQUENCE 21 AA; 2235 MW; 30C8B7B7E0D8F244 CRC64;

Query Match 22.9%; Score 32; DB 9; Length 21;
Best Local Similarity 62.5%; Pred. No. 4.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 SIVDRPDF 14
  | : | | | |
Db 3 SMADKPDF 10

RESULT 8
ID Q9QV92 PRELIMINARY; PRT; 22 AA.
AC Q9QV92;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE P93 CALCIUM-BINDING PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101118;
RN [1]
RP SEQUENCE
RX MEDLINE=93179436; PubMed=8440713;
RA Glichrist J.S., Pierce G.N.;
RT "Identification and purification of a calcium-binding protein in
RT hepatic nuclear membranes.";
RL J. Biol. Chem. 268:4291-4299(1993).

```

```

DR InterPro: IP8001580; Calreticulin.
DR Pfam: PF00262; calreticulin; 1.
DR ProDom: PD001866; Calreticulin; 1.
SQ SEQUENCE 22 AA; 2407 MW; B82830E4477E25F9 CRC64;

Query Match 22.9%; Score 32; DB 11; Length 22;
Best Local Similarity 28.6%; Pred. No. 4.5e+02;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYGYQSVDRPDF 14
  || : : | | :
Db 7 CGVWRPMDPNPY 20

RESULT 9
ID Q85496 PRELIMINARY; PRT; 23 AA.
AC Q85496;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Bovine leukemia virus (BLV).
OC Viruses; Retrovirus; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11901;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87198886; PubMed=3033284;
RA Rice N.R., Simek S.L., Dubois G.C., Showalter S.D., Gilden R.V.,
RA Stephens R.M.;
RT "Expression of the bovine leukemia virus X region in virus-infected
RT cells.";
RL J. Virol. 61:1577-1585(1987).
DR EMBL; M16017; AAA87336.1; -.
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2877 MW; F85E7CBED5440B08 CRC64;

Query Match 22.9%; Score 32; DB 15; Length 23;
Best Local Similarity 50.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVRSITL 24
  | : | | | | : | |
Db 11 PQPIRWVSL 20

RESULT 10
ID P82340 PRELIMINARY; PRT; 14 AA.
AC P82340;
DT 01-JUN-2000 (TREMBlrel. 14, Created)
DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF THYLAKOID LUMEN (SPOT204)
DE (FRAGMENT).
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.
RC STRAIN=CV. DE GRACE; TISSUE=LEAF;
RX MEDLINE=20181728; PubMed=10715320;
RA Peltier J.-B., Friso G., Kalume D.E., Roepstorff P., Nilsson F.,
RA Adamska I., van Wijk K.J.;
RT "Proteomics of the chloroplast: systematic identification and
RT targeting analysis of luminal and peripheral thylakoid proteins.";
RL Plant Cell 12:319-341(2000).
CC -|- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN.
CC -|- DEVELOPMENTAL STAGE: UNFOLDED AND FULLY DEVELOPED LEAVES.

```

KW Chloroplast; Thylakoid membrane.

SQ NON_TER 14 14
FT SEQUENCE 14 AA; 1505 MW; 2EABFAF980F3D7D7 CRC64;Query Match 22.1%; Score 31; DB 10; Length 14;
Best Local Similarity 52.5%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYQSIQVDR 11

| : | : |||

Db 4 GFQPVVDR 11

RESULT 11

Q56610 ID Q56610 PRELIMINARY; PRT; 18 AA.

AC Q56610;

DT 01-NOV-1996 (TReMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)

DE 01-NOV-1998 (TReMBLrel. 08, Last annotation update)

DE ACCA (FRAGMENT).

GN ACCA.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C6706;

RX MEDLINE=97074686; PubMed=8917113;

RA Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,

RA Morris J.G.;

RT "cloning and characterization of dnaE, encoding the catalytic subunit

of replicative DNA polymerase III, from Vibrio cholerae strain

RT C6706."

RL Gene 175:281-283(1996).

DR EMBL; U30472; AAC44579.1; -.

FT NON_TER 18 18

SQ SEQUENCE 18 AA; 2153 MW; 18EBDAD212842EF CRC64;

Query Match

Best Local Similarity 22.1%; Score 31; DB 2; Length 18;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 DFKPPIV 19

| : | : |||

Db 7 DFEKPIV 13

RESULT 12

Q95JA2

ID Q95JA2 PRELIMINARY; PRT; 18 AA.

AC Q95JA2;

DT 01-DEC-2001 (TReMBLrel. 19, Created)

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)

DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)

DE INTEGRIN ALPHA 2 SUBUNIT (FRAGMENT).

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=CARDIOVASCULAR;

RA Markmann A., Kresse H.;

RT "Regulation of VSMC Differentiation.";

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF330200; AAL09466.1; -.

FT NON_TER 1 1

FT NON_TER 18 18

SQ SEQUENCE 18 AA; 1944 MW; DFEAC6E14A5B0EF CRC64;

Query Match 22.1%; Score 31; DB 6; Length 18;
Best Local Similarity 45.5%; Pred. No. 5.2e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 PKPIVRSITLC 25

| : | : ||

Db 4 PSDVNSLDLC 14

RESULT 13

Q9X3H0

ID Q9X3H0 PRELIMINARY; PRT; 21 AA.

AC Q9X3H0;

DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).

GN PETD.

OS Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI_TaxID=1220;

RN [1]

RP SEQUENCE FROM N.A.

RA Urbach E., Chisholm S.W.;

RT "Genetic diversity in Prochlorococcus populations flow cytometrically

sorted from the Sargasso Sea and Gulf Stream.";

RL Limnol. Oceanog. 43:1615-1630(1998).

DR EMBL; AF070163; AAD23212.1; -.

FT NON_TER 21 21

SQ SEQUENCE 21 AA; 2280 MW; CC946C4A2034E54F CRC64;

Query Match

Best Local Similarity 22.1%; Score 31; DB 2; Length 21;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 7 SIIVDRPDPKPIVRS 21

| : | : ||

Db 2 SVLKKPLADPKLRA 16

RESULT 14

Q9X3J8

ID Q9X3J8 PRELIMINARY; PRT; 21 AA.

AC Q9X3J8;

DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)

DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).

GN PETD.

OS Prochlorococcus sp.

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI_TaxID=1220;

RN [1]

RP SEQUENCE FROM N.A.

RA Urbach E., Chisholm S.W.;

RT "Genetic diversity in Prochlorococcus populations flow cytometrically

sorted from the Sargasso Sea and Gulf Stream.";

RL Limnol. Oceanog. 43:1615-1630(1998).

DR EMBL; AF070190; AAD23228.1; -.

FT NON_TER 21 21

SQ SEQUENCE 21 AA; 2298 MW; CC946FFCC02C854F CRC64;

Query Match

Best Local Similarity 21.4%; Score 30; DB 2; Length 21;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 7 SIIVDRPDPKPIVRS 21

| : | : ||

Db 2 STLKKPLADPKLRS 16

RESULT 15

Q9BM10
ID Q9BM10 PRELIMINARY; PRT; 24 AA.
AC Q9BM10;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GYPSY-LIKE REVERSE TRANSCRIPTASE (FRAGMENT).
OS Spongilla lacustris (Freshwater sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC Haplosclerida; Spongillidae; Spongilla.
OX NCBI_TaxID=6055;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=GRT-G6 RETROTRANSPOSON;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR ENBL; AY013996; AAC59968.1; -;
KW RNA-directed DNA polymerase.
FT NON_TER 1 1
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2499 MW; 4A03D8E3E7A4D12 CRC64;

Query Match 21.1%; Score 29.5; DB 5; Length 24;
Best Local Similarity 54.5%; Pred. No. 1.2e+03;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 2 GYGQSIIVDRP 12
|| :| :| :|
Db 6 GY-FQEIMDQP 15

Search completed: August 26, 2002, 10:29:35
Job time: 650 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:18:42 ; Search time 79.43 Seconds
(without alignments)
34.960 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 GCGYQSVDRPDPKPIVRSLITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 253524

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

```
1: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneq-emb1/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	25	AA1980001	Optimised IgE-CH3
2	106	75.7	25	AA1980000	Optimised IgE-CH3
3	74	52.9	25	AA1980077	Optimised IgE-CH3
4	66	47.1	25	AA1799999	Optimised IgE-CH3
5	64	45.7	25	AA1799998	Optimised IgE-CH3
6	64	45.7	25	AA191212	Modified human IgE
7	64	45.7	25	AA191212	Peptide sequence o
8	54	38.6	15	AA1968602	Antibody 15A.2 mur
9	53	37.9	17	AA150899	Antibody 15A.2 swi
10	45	32.1	17	AA192179	Peptide which bind
11	41	29.3	17	AA150893	Antibody 15A.2 bin

```
12 41 29.3 17 21 AAY50894
13 41 29.3 17 21 AAY50897
14 41 29.3 20 18 AAW24102
15 38.5 27.5 16 17 AAW09376
16 38.5 27.5 18 17 AAW09378
17 38 27.1 12 22 ABB52359
18 38 27.1 12 22 AAU25118
19 38 27.1 12 22 AAU15462
20 38 27.1 16 14 AAR30832
21 36 25.7 20 17 AAW05484
22 36 25.7 20 17 AAW05416
23 36 25.7 20 17 AAW16970
24 36 25.7 20 18 AAW25345
25 35.5 25.4 16 17 AAW09377
26 35 25.0 17 21 AAB20671
27 35 25.0 22 14 AAR41158
28 33.5 23.9 11 17 AAW17451
29 33 23.6 17 15 AAR51405
30 33 23.6 17 21 AAY50896
31 32 22.9 14 19 AAW59655
32 32 22.9 14 22 AAB88089
33 32 22.9 15 21 AAY98820
34 32 22.9 15 21 AAY98912
35 32 22.9 18 17 AAW05469
36 32 22.9 18 18 AAW37677
37 32 22.9 20 14 AAR41190
38 32 22.9 20 14 AAR41191
39 32 22.9 20 14 AAR41192
40 32 22.9 20 18 AAW10148
41 32 22.9 20 18 AAW10149
42 32 22.9 21 21 AAY58800
43 32 22.9 22 14 AAR41118
44 32 22.9 22 14 AAR41119
45 32 22.9 23 16 AAR70227
```

ALIGNMENTS

```
RESULT 1
AA1980001
ID AAY80001 standard; Peptide; 25 AA.
AC AAY80001;
XX
XX
XX 15-MAY-2000 (first entry)
XX
XX Optimised IgE-CH3 domain antigen peptide for mouse IgE.
```

```
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
```

```
XX Mus sp.
XX Synthetic.
```

```
XX WO9967293-A1.
```

```
XX 29-DEC-1999.
```

```
XX 21-JUN-1999; 99WO-US13959.
```

```
XX 20-JUN-1998; 98US-0100287.
```

```
XX (UNBI-) UNITED BIOMEDICAL INC.
```

```
XX Wang CY, Walfield AM;
```

```
XX WPI; 2000-160578/14.
```

```
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy
```

CC	specific for a target effector site on the epsilon-heavy chain of IgE.
CC	and so preventing triggering and activation of mast cells and basophils
CC	and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC	containing (I) are used for active immunisation against IgE-mediated
CC	allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC	dermatitis. Nucleic acids that encode these compounds are useful for
CC	recombinant production of corresponding peptides or in DNA vaccines.
CC	conjugates of (I) that include a promiscuous T helper cell epitope
CC	(functional in genetically diverse subjects), in addition to a B cell
CC	target epitope, have increased immunogenicity and may include cyclic
CC	constraints (disulfide bridge) to stabilise conformational features and
CC	maximize cross-reactivity to the natural target. They induce safe
CC	(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC	acid sequences used in the exemplification of the present invention.
XX	
SQ	Sequence 25 AA;
	Query Match 75.7%; Score 106; DB 21; Length 25;
	Best Local Similarity 80.0%; Pred. No. 1.6e-09;
	Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy	1 CGYGYQSIIVDRDPFKPIVRSITLC 25
Db	1 cgegyqsrvdhpfkpivrstkc 25
RESULT	3
AAAY80077	AAAY80077 standard; Peptide; 25 AA.
ID	AAAY80077
AC	AAAY80077;
XX	
DT	15-MAY-2000 (first entry)
DE	Optimised IgE-CH3 domain antigen peptide for horse IgE.
XX	
KW	Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW	immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW	antibody; allergy; allergic disease; immunisation; anti-allergic;
KW	anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX	
OS	Equus caballus.
OS	Synthetic.
XX	
PN	WO9967293-A1.
PD	29-DEC-1999.
XX	
PF	21-JUN-1999; 99WO-US13959.
XX	
PR	20-JUN-1998; 98US-0100287.
XX	
PA	(UNBI-) UNITED BIOMEDICAL INC.
XX	
PI	Wang CY, Walfield AM;
XX	
DR	WPI; 2000-160578/14.
XX	
PT	New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT	for immunization against allergy -
XX	
PS	Claim 1; Page 146; 155pp; English.
XX	
CC	The present invention describes immunoglobulin E (IgE)-CH3 domain
CC	antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC	and anti-asthmatic properties. (I) induces polyclonal antibodies
CC	specific for a target effector site on the epsilon-heavy chain of IgE,
CC	and so preventing triggering and activation of mast cells and basophils
CC	and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC	containing (I) are used for active immunisation against IgE-mediated
CC	allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC	dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 52.9%; Score 74; DB 21; Length 25;
 Best Local Similarity 52.0%; Pred. No. 0.00015;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGYGQSIIVDRDPFKPIVRSITLC 25
 II I I I I I I I I I I
 Db 1 CGETKSTVSHDLPREVRSIAK 25

RESULT 4
 AAY79999
 ID AAY79999 standard; Peptide; 25 AA.
 XX
 AC AAY79999;

DT 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Canis sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 PS Claim 1; Page 99; 155pp; English.

XX The present invention describes Immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 47.1%; Score 66; DB 21; Length 25;
 Best Local Similarity 56.0%; Pred. No. 0.0027;
 Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGQSIIVDRDPFKPIVRSITLC 25
 II I I I I I I I I I I
 Db 1 CGETYSRVTHPLPKDIVRSIAK 25

RESULT 5
 AAY79998
 ID AAY79998 standard; Peptide; 25 AA.
 XX
 AC AAY79998;

DT 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Homo sapiens.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 PS Claim 1; Page 21; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

```

Query Match      45.7%; Score 64; DB 21; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0055;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYQYSIVDRPDFPKPIVRSLTLC 25
   || ||| | | | : || | |
Db 1 cgetyqsrvtphlpralmrstkc 25

RESULT 6
AA91212
ID AA91212 standard; peptide: 25 AA.
AC AA91212;
XX
DT 22-MAY-2000 (first entry)
DE Modified human IgE CH3 domain, SEQ ID NO:92.
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CTFP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO9966957-A2.
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
PI Wang CY;
XX
DR WPI: 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus
XX
PS Example 6; Page 40; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CTFP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer; prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AA91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AA91122-Y91142,
CC AA91226 and AA91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AA91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences

```

```

CC AA91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AA91156-Y91196, AA91227 and AA91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AA91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AA91200 is somatostatin, and AA91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AA91208 is a
CC human CD4 CDR2-like domain antigenic site, and AA91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AA90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AA90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AA91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AA91221-Y91222 comprise this peptide and a Th
CC epitope. AA91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AA91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AA91228-Y91231 represent
CC CTFP-derived peptides and AA91232-Y91241 are immunogens comprising a
CC CTFP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AA91247 and AA91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AA91248-Y91251 and
CC AA91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AA91198 and AA91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX
SQ Sequence 25 AA;

```

```

Query Match      45.7%; Score 64; DB 21; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0055;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYQYSIVDRPDFPKPIVRSLTLC 25
   || ||| | | | : || | |
Db 1 cgetyqsrvtphlpralmrstkc 25

```

```

RESULT 7
AA91212
ID AA91212 standard; peptide: 25 AA.
XX
AC AA91212;
XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide sequence of the invention.
XX
KW Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX
OS Unidentified.
XX
PN WO9966952-A1.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13960.
XX
PR 20-JUN-1998; 98US-0100414.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
PI Wang CY;
XX
DR WPI: 2000-160562/14.
XX
PT New peptide immunogen containing luteinising hormone-releasing hormone

```

PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer -
 XX
 XX
 PS Disclosure; Page 92; 102pp; English.
 XX
 CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.
 XX
 SQ Sequence 25 AA;

Query Match 45.7%; Score 64; DB 21; Length 25;
 Best Local Similarity 48.0%; Pred. No. 0.0055;
 Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGVSVDRPDPFKPIVRSI 25
 || ||| | | : ||| |
 Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 8
 AAY50899
 ID AAY50899 standard; peptide; 15 AA.

XX
 AC AAY50899;
 XX
 DT 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 murine IgE binding epitope 1.
 XX
 KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.

PF 09-APR-1999; 99EP-0107035.
 XX
 PR 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.

XX (IDEX-) IDEXX LAB INC.

PI Lawton R, Mermer B, Francoeur G;
 XX
 DR WPI; 2000-040833/04.

PT Binding proteins used for treatment or prophylaxis of canine allergy -
 XX
 PS Disclosure; Fig 7; 30pp; English.

XX This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which

CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 SQ Sequence 15 AA;

Query Match 38.6%; Score 54; DB 21; Length 15;
 Best Local Similarity 76.5%; Pred. No. 0.11;
 Matches 13; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 6 QSIIVDRPDPFKPIVRSI 22
 | ||| ||| |||||
 Db 1 qcivdhpdf--plvrsi 15

RESULT 9
 AAY50898
 ID AAY50898 standard; peptide; 17 AA.

XX
 AC AAY50898;

XX
 DT 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 swine IgE binding epitope 1.

XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.

XX
 PN EP957111-A2.

XX
 PD 17-NOV-1999.

XX
 PF 09-APR-1999; 99EP-0107035.

XX
 PR 09-APR-1998; 98US-0058331.

XX
 PR 30-MAR-1999; 99US-0281760.

XX (IDEX-) IDEXX LAB INC.

XX
 PI Lawton R, Mermer B, Francoeur G;
 XX
 DR WPI; 2000-040833/04.

XX Binding proteins used for treatment or prophylaxis of canine allergy -
 XX
 PS Disclosure; Fig 7; 30pp; English.

XX This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.

XX
 SQ Sequence 17 AA;

Query Match 37.9%; Score 53; DB 21; Length 17;
 Best Local Similarity 71.4%; Pred. No. 0.18;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 VDRPDPFKPIVRSI 22
 | || |||||
 Db 4 vthpdlpkplrsi 17

```

RESULT 10
AAR02179
ID AAR02179 standard; protein; 17 AA.
XX
AC AAR02179;
XX
DT 02-AUG-1990 (first entry)
XX
DE Peptide which binds to Immunoglobulin E (IgE) receptors inhibiting IgE
DE binding.
XX
KW Immunoglobulin E (IgE); anti-allergic drug; IgE binding inhibitor.
XX
FH Key Location/Qualifiers
FT Misc-difference 1
FT /label=OTHER
FT /note="H-Pro"
FT Misc-difference 17
FT /label=OTHER
FT /note="Ala-OH"
XX
PN JP01299298-A.
XX
PD 04-DEC-1989.
XX
PF 27-MAY-1988; 88JP-0129721.
XX
PR 27-MAY-1988; 88JP-0129721.
XX
PA (KYOW ) KYOWA HAKKO KOGYO KK.
XX
DR WPI; 1990-019330/03.
XX
PT Peptides which bind to IgE receptors inhibiting IgE binding -
PT used as anti-allergic drug having no side effects
XX
PS Claim 2; Page 843; 8pp; Japanese.
XX
CC It is the new sequence No 11 and has an anti-allergic effect. It binds to
CC IgE receptors on cell surfaces and inhibits IgE binding to receptors. It
CC can be used as an anti-allergic drug, having no side effects.
XX
SQ Sequence 17 AA;

Query Match 32.1%; Score 45; DB 11; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PKPIVRSIT 23
Db 1 pkplvrsit 9

RESULT 11
AAV50893
ID AAV50893 standard; peptide; 17 AA.
XX
AC AAV50893;
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 binding peptide 10 from PhdC7c phage display library.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP57111-A2.
XX
PD 17-NOV-1999.
XX
PS Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically

```

```

PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
XX Disclosure; Fig 6; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 17 AA;

Query Match 29.3%; Score 41; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 VDRPDPFKPIVRSI 22
Db 4 vchphlpkdivrsi 17

RESULT 12
AAV50894
ID AAV50894 standard; peptide; 17 AA.
XX
AC AAV50894;
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 canine IgE binding epitope 1.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically

```

CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 SQ Sequence 17 AA;

Query Match 29.3%; Score 41; DB 21; Length 17;
 Best Local Similarity 64.3%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 VDRPDPFKPIVRSI 22
 | | | | |
 Db 4 vtphlpkdivrsi 17

RESULT 13
 AAY50897
 ID AAY50897 standard; peptide; 17 AA.
 XX
 AC AAY50897;
 XX

DT 24-FEB-2000 (first entry)
 XX

DE Antibody 15A.2 feline IgE binding epitope 1.
 XX

KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.
 XX

PF 09-APR-1999; 99EP-0107035.
 XX

PR 09-APR-1998; 98US-0058331.
 XX

PR 30-MAR-1999; 99US-0281760.
 XX

PA (IDEX-) IDEXX LAB INC.
 XX

PI Lawton R, Mermer B, Francoeur G;
 XX

DR WPI; 2000-040833/04.
 XX

PT Binding proteins used for treatment or prophylaxis of canine allergy -
 PS Disclosure; Fig 7; 30pp; English.
 XX

CC This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 SQ Sequence 17 AA;

Query Match 29.3%; Score 41; DB 21; Length 17;
 Best Local Similarity 58.8%; Pred. No. 14;
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 QSIIVDRPDKPIVRSI 22
 | | | | |
 Db 1 qckvthpdlplvirs1 17

RESULT 14
 AAW24102
 ID AAW24102 standard; peptide; 20 AA.
 XX
 AC AAW24102;
 XX

DT 21-NOV-1997 (first entry)
 XX

DE Canine immunoglobulin E peptide 5.
 XX

KW Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
 OS Canis familiaris.
 XX
 PN JP09169795-A.
 XX

PD 30-JUN-1997.
 XX

PF 22-DEC-1995; 95JP-0334381.
 XX

PR 22-DEC-1995; 95JP-0334381.
 XX

PA (HITB) HITACHI CHEM CO LTD.
 XX

DR WPI; 1997-389423/36.
 DR N-PSDB; AAT85651.
 XX

PT Canine immunoglobulin E peptide fragment and related DNA - useful
 PT for the preparation of anti-canine immunoglobulin E antibody
 XX
 PS Claim 2; Page 9; 12pp; Japanese.
 XX

CC AAW24098-106 are peptide fragments containing at least 5 continuous
 CC amino acids of the partial canine immunoglobulin E (IgE) protein shown
 CC in AAW24097. The peptides are used for the preparation of anti-canine
 CC IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
 CC of canine allergies.
 XX
 SQ Sequence 20 AA;

Query Match 29.3%; Score 41; DB 18; Length 20;
 Best Local Similarity 64.3%; Pred. No. 16;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 VDRPDPFKPIVRSI 22
 | | | | |
 Db 4 vtphlpkdivrsi 17

RESULT 15
 AAW09376
 ID AAW09376 standard; peptide; 16 AA.
 XX
 AC AAW09376;
 XX

DT 17-JUN-1997 (first entry)
 XX

DE Proline-rich antibacterial peptide from *Podisus maculiventris*.
 XX
 KW *Podisus maculiventris*; proline-rich; pathogenic; bacteria; bug;
 KW Gram negative; plant treatment.
 XX
 OS *Podisus maculiventris*.
 XX
 PN FR2732345-A1.
 XX

PD 04-OCT-1996.
 XX
 PF 03-APR-1995; 95FR-0004130.
 XX
 PR 03-APR-1995; 95FR-0004130.
 XX
 PA (RHON) RHONE POULENC AGROCHIMIE.
 XX
 PI Bulet P, Hetru C, Hoffman J, Tchernych S;
 XX
 DR WPI; 1996-457711/46.
 XX
 PT Proline-rich antibacterial peptide(s) from Podisus maculiventris -
 PT esp. useful to protect plants against pathogenic bacteria, also for
 PT antibacterial therapy of animals
 XX
 PS Claim 2; Page 14; 16pp; English.
 XX
 CC AAM09375-W09378 are specific examples of antibacterial peptides from the
 CC bug Podisus maculiventris. The proline-rich peptides have antibacterial
 CC activity against gram negative bacteria. They are useful for protecting
 CC plants against pathogenic bacteria, but could also be used for
 CC antibacterial therapy of animals, including humans. A generic peptide
 CC having the sequence VDKPDYRPRPX is also given in the specification,
 CC where X is a peptide comprising at least one PRP motif (see AAM17451).
 XX
 SQ Sequence 16 AA;

Query Match 27.5%; Score 38.5; DB 17; Length 16;
 Best Local Similarity 43.8%; Pred. NO. 31;
 Matches 7; Conservative 4; Mismatches 0; Indels 5; Gaps 1;
 OY 9 VDRPDF-----PKPIV 19
 ||:||: ||:||:
 Db 1 vdkpdyrprgprpi 16

Search completed: August 26, 2002, 10:18:42
 Job time: 698 sec

us-09-701-623c-8.closed.ra1

Mon Aug 26 11:09:36 2002

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:19:37 ; Search time 33.13 Seconds
(without alignments)
18.432 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGGYQSVDRPDPFKPIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues 134663

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	64	45.7	25	3	US-09-100-414B-95
2	64	45.7	25	4	US-09-303-323-95
3	38	27.1	16	1	US-08-447-010-14
4	38	27.1	21	1	US-08-447-411-19
5	38	27.1	21	1	US-08-447-411-59
6	38	27.1	21	2	US-08-662-227-16
7	38	27.1	21	4	US-09-017-947-16
8	38	27.1	22	1	US-08-447-411-72
9	38	27.1	22	2	US-08-662-227-30
10	38	27.1	22	4	US-09-017-947-30
11	36	25.7	20	4	US-08-602-999A-120
12	36	25.7	20	4	US-08-630-915A-180
13	36	25.7	20	4	US-08-630-915A-222
14	35	25.0	21	1	US-08-447-411-58
15	35	25.0	21	2	US-08-662-227-15
16	35	25.0	21	4	US-09-017-947-15
17	35	25.0	22	1	US-08-447-411-71
18	35	25.0	22	2	US-08-146-028-101
19	35	25.0	22	2	US-08-662-227-29
20	35	25.0	22	4	US-08-723-425A-101
21	35	25.0	22	4	US-09-112-206-101
22	35	25.0	22	4	US-09-017-947-29
23	32	22.9	13	6	5177197-5
24	32	22.9	18	3	US-08-630-916A-109
25	32	22.9	18	4	US-08-630-915A-165
26	32	22.9	20	2	US-08-466-975A-18
27	32	22.9	20	2	US-08-466-975A-19

28	32	22.9	20	2	US-08-391-671A-18
29	32	22.9	20	2	US-08-391-671A-19
30	32	22.9	20	3	US-08-467-902A-18
31	32	22.9	20	3	US-08-467-902A-19
32	32	22.9	20	4	US-09-275-265-18
33	32	22.9	20	4	US-09-275-265-19
34	32	22.9	21	2	US-08-448-600-9
35	32	22.9	21	3	US-08-937-610-17
36	32	22.9	21	4	US-09-122-127D-6
37	32	22.9	22	2	US-08-146-028-61
38	32	22.9	22	2	US-08-146-028-122
39	32	22.9	22	4	US-08-723-425A-61
40	32	22.9	22	4	US-08-723-425A-62
41	32	22.9	22	4	US-08-723-425A-122
42	32	22.9	22	4	US-09-112-206-61
43	32	22.9	22	4	US-09-112-206-62
44	32	22.9	22	4	US-09-112-206-122
45	32	22.9	22	4	US-09-112-206-122

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-95

Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
QY 1 CGGYQSVDRPDPFKPIVRSITLC 25
DB 1 CGGYQSVDRPDPFKPIVRSITLC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Meng, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 45.78; Score 64; DB 4; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0021;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
QY 1 CGYGYQSIYDRDPKPIVRSITLC 25
Db 1 CGETYSRVTHPLPRALMRSTTKC 25

RESULT 3
US-08-447-010-14
; Sequence 14, Application US/08447010
; Patent No. 5770718
; GENERAL INFORMATION:
; APPLICANT: MOFFATT, BARBARA
; TITLE OF INVENTION: GENE FOR APT FROM PLANT TISSUE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurrey
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,010
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/230,695
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,132
; FILING DATE: 26-MAY-1992
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1811-183 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-447-010-14
Query Match 27.1%; Score 38; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 PDFPKP 17
Db 3 PDFPKP 8
RESULT 4
US-08-447-411-19
; Sequence 19, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVE1, AND CVP2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Naja naja
US-08-447-411-19

Query Match 27.1%; Score 38; DB 1; Length 21;
Best Local Similarity 42.9%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGYQSVIVDRDPFK 16
DB 8 FGDDNIISRSDFPE 21

RESULT 5

US-08-447-411-59
Sequence 59, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHORST
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Naja naja
US-08-447-411-59

Query Match 27.1%; Score 38; DB 1; Length 21;
Best Local Similarity 42.9%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 3 YGYQSVIVDRDPFK 16
DB 8 FGDDNIISRSDFPE 21
RESULT 6
US-08-662-227-16
Sequence 16, Application US/08662227
Patent No. 5923320
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-662-227-16

Query Match 27.1%; Score 38; DB 2; Length 21;
Best Local Similarity 42.9%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGYQSVIVDRDPFK 16
DB 8 FGDDNIISRSDFPE 21

RESULT 7

US-09-017-947-16
Sequence 16, Application US/09017947
Patent No. 6303754
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF

```
;
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 16:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; US-09-017-947-16

Query Match 27.1%; Score 38; DB 4; Length 21;
Best Local Similarity 42.9%; Pred. No. 16;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGYQSVIVDRDPKP 16
Db 8 FGDDNIISRDPE 21

RESULT 8
US-08-447-411-72
; Sequence 72, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
```

```
;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Naja naja
; US-08-447-411-72

Query Match 27.1%; Score 38; DB 1; Length 22;
Best Local Similarity 42.9%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGYQSVIVDRDPKP 16
Db 8 FGDDNIISRDPE 21

RESULT 9
US-08-662-227-30
; Sequence 30, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

MOLECULE TYPE: peptide
US-08-662-227-30

Query Match 27.1%; Score 38; DB 2; Length 22;
Best Local Similarity 42.9%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGQSVDRDPFK 16
DB 8 FGDDNIISRDFFE 21

RESULT 10

US-09-017-947-30
; Sequence 30, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017, 947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-017-947-30

Query Match 27.1%; Score 38; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 17;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGQSVDRDPFK 16
DB 8 FGDDNIISRDFFE 21

RESULT 11

US-08-602-999A-120
; Sequence 120, Application US/08602999A

Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: O'ILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-120

Query Match 25.7%; Score 36; DB 4; Length 20;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVRSITL 24
DB 9 PKPIRSVSL 18

RESULT 12

US-08-630-915A-180
; Sequence 180, Application US/08630915A
; Patent No. 6309820
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: HOFFMAN, No. 6309820h
; APPLICANT: KAY, Brian K.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: MCCONNELL, Stephen J.
; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York

COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
INFORMATION FOR SEQ ID NO: 180:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-630-915A-180

Query Match 25.7%; Score 36; DB 4; Length 20;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 15 PKPIVRSITL 24
||| :||::|
Db 9 PKPIRSVSL 18

RESULT 13
US-08-630-915A-222
Sequence 222, Application US/08630915A
Patent No. 6309820
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: HOFFMAN, No. 6309820h
APPLICANT: KAY, Brian K.
APPLICANT: FOWLKES, Dana M.
APPLICANT: MCCONNELL, Stephen J.
TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL
TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND
TITLE OF INVENTION: USING SAME
NUMBER OF SEQUENCES: 227
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,915A
FILING DATE: 03-APR-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-174
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 222:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Other
OTHER INFORMATION: May or may not have carboxy-terminal
OTHER INFORMATION: amide and/or biotinylated N-terminal
US-08-630-915A-222

Query Match 25.7%; Score 36; DB 4; Length 20;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 15 PKPIVRSITL 24
||| :||::|
Db 9 PKPIRSVSL 18

RESULT 14
US-08-447-411-58
Sequence 58, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-447-411-58

Query Match 25.0%; Score 35; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 IVDPRDPFK 16
I: | | | | |
Db 13 IISRSDFPK 21

RESULT 15

US-08-662-227-15
; Sequence 15, Application US/08662227
; Patent No. 592320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-662-227-15

Query Match 25.0%; Score 35; DB 2; Length 21;
Best Local Similarity 66.7%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 IVDPRDPFK 16
I: | | | | |
Db 13 IISRSDFPK 21

Search completed: August 26, 2002, 10:19:37
Job time: 393 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:20:42 ; Search time 43.94 Seconds
(without alignments)
54.671 Million cell updates/sec

Title: US-09-701-623c-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 4981

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	35	25.9	24	2 S29749	serum albumin - do
2	31	23.0	19	2 PS0186	superoxide dismuta
3	30	22.2	16	2 S30384	hypothetical prote
4	29	21.5	23	2 I28027	superoxide dismuta
5	29	21.5	25	2 JP0045	ribosomal protein
6	28	20.7	20	2 A61093	glue protein - Cal
7	27.5	20.4	15	2 PS0251	15K protein 5106 -
8	27	20.0	11	2 S78765	ribosomal protein
9	27	20.0	23	2 S41390	p7 protein - human
10	26.5	19.6	25	2 A58647	alphaA-conotoxin p
11	26	19.3	9	2 I46016	cytokeratin 4 - bo
12	26	19.3	15	4 I38335	hypothetical TEL/M
13	26	19.3	16	2 A49255	T-cell receptor be
14	26	19.3	19	2 PA0012	superoxide dismuta
15	26	19.3	23	2 A41263	kinase-related tra
16	26	19.3	24	2 S35641	DNA (cytosine-5)-
17	25	18.5	10	2 H37196	bradykinin-potenti
18	25	18.5	15	2 PA0071	superoxide dismuta
19	25	18.5	17	2 PH1822	T cell receptor al
20	25	18.5	20	2 D49164	chromogranin-B - r
21	25	18.5	21	2 S47207	T-cell receptor J-
22	25	18.5	23	2 F61491	seed protein ws-6
23	24.5	18.1	24	2 S42785	relaxin - baboon (
24	24	17.8	10	2 B37196	bradykinin-potenti
25	24	17.8	14	2 S48685	extension protein
26	24	17.8	18	2 PC2280	metalloprotein
27	24	17.8	18	2 PC2280	prolylendopeptidas
28	24	17.8	20	2 S23981	outer layer protei
29	24	17.8	20	2 A61276	superoxide dismuta

30 24 17.8 23 2 S23637 hypothetical prote
31 24 17.8 25 2 JP0064 ribosomal protein
32 23 17.0 13 2 A54326 glandular kallikre
33 23 17.0 14 1 NTKN1M alpha-conotoxin MI
34 23 17.0 14 2 S74128 superoxide dismuta
35 23 17.0 14 2 PH1806 T cell receptor al
36 23 17.0 15 2 PA0009 seed storage prote
37 23 17.0 15 2 PA0088 protein QF200051 -
38 23 17.0 17 2 PH1802 T cell receptor al
39 23 17.0 18 2 S18386 110K protein - mou
40 23 17.0 19 2 PQ0409 RNA-directed RNA p
41 23 17.0 20 2 S28435 major outer membra
42 23 17.0 20 2 T01691 hypothetical prote
43 23 17.0 20 2 A05310 apolipoprotein E -
44 23 17.0 21 2 A56901 nerve growth facto
45 23 17.0 23 2 S72535 probable acr-2 reg

ALIGNMENTS

RESULT 1

S29749
serum albumin - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S29749
R:Dixon, J.W.; Sarkar, B.
J. Biol. Chem. 249, 5872-5877, 1974
A:Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide
A:Reference number: S29749; MUID:75011422
A:Accession: S29749
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <DIX>
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match 25.9%; Score 35; DB 2; Length 24;
Best Local Similarity 40.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 3 ETYKSTVSH--PDLPREVRSI 22
| | | | | : | | | | | : | :
Db 1 EAYKSEIAHRYNDLGEHFRGL 22

RESULT 2

PS0186
superoxide dismutase (EC 1.15.1.1) (Mn) - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 05-Mar-1999
C:Accession: PS0186
R:Kamo, M.; Tsugita, A.
submitted to JIPID, June 1991
A:Reference number: PS0184
A:Accession: PS0186
A:Molecule type: protein
A:Residues: 1-19 <KAM>
A:Experimental source: callus
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg
C:Superfamily: superoxide dismutase (Mn)
C:Keywords: manganese; metalloprotein; oxidoreductase

Query Match 23.0%; Score 31; DB 2; Length 19;
Best Local Similarity 75.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 TVSHPDLP 15
| | | | |
Db 3 TVALPDLP 10

RESULT 3
S30384
hypothetical protein - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 08-Oct-1999
C:Accession: S30384
R:Willey, D.L.; Caswell, D.A.; Lowe, C.R.; Bruce, N.C.
Biochem. J. 290, 539-544, 1993
A:Title: Nucleotide sequence and over-expression of morphine dehydrogenase, a plasmid-encoded protein from *Pseudomonas putida*
A:Reference number: S30384
A:Accession: S30384
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-16 <WIL>
A:Cross-references: EMBL:M94775; NID:gl621506; PIDN:AAB17357.1; PID:gl51367

Query Match 22.2%; Score 30; DB 2; Length 16;
Best Local Similarity 53.8%; Pred. No. 4.4e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 13 DLPREVRSIAKC 25
||||| ||| |
Db 5 DLPFRS--RSLRSC 15

RESULT 4
I28027
superoxide dismutase (EC 1.15.1.1) (Mn) - curled-leaved tobacco (fragment)
C:Species: Nicotiana glauca (curled-leaved tobacco)
C:Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 05-Mar-1999
C:Accession: I28027
R:Bauw, G.; De Loose, M.; Inze, D.; Van Montagu, M.; Vandekerckhove, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 4806-4810, 1987
A:Title: Alterations in the phenotype of plant cells studied by NH2-terminal amino acid sequencing of superoxide dismutase (Mn)
A:Reference number: A94167
A:Accession: I28027
A:Molecule type: protein
A:Residues: 1-23 <BAU>
C:Function: superoxide dismutase
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase (Mn)
C:Keywords: metalloprotein; mitochondrion; oxidoreductase

Query Match 21.5%; Score 29; DB 2; Length 23;
Best Local Similarity 75.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TVSHPDLP 15
| | | | |
Db 3 TVSLPDLP 10

RESULT 5
JP0045
ribosomal protein L30 - Staphylococcus aureus (fragment)
C:Species: Staphylococcus aureus
C:Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 03-May-1996
C:Accession: JP0045
R:Ochi, K.
submitted to JPIID, February 1994
A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal protein L30
A:Reference number: JP0042
A:Accession: JP0045
A:Molecule type: protein
A:Residues: 1-25 <OCH>
C:Superfamily: Escherichia coli ribosomal protein L30
C:Keywords: protein biosynthesis; ribosome

Query Match 21.5%; Score 29; DB 2; Length 25;
Best Local Similarity 23.5%; Pred. No. 9.9e+02;
Matches 4; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

QY 6 KSTVSHPDLPREVRSI 22
| | | | |
Db 9 RSVIGRPETQRKTVEAL 25

RESULT 6
A61093
glue protein - California mussel (fragments)
N:Alternate names: adhesive polyphenolic protein
C:Species: Mytilus californianus (California mussel)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: A61093
R:Waite, J.H.
J. Comp. Physiol. B 156, 491-496, 1986
A:Title: Mussel glue from Mytilus californianus Conrad: a comparative study.
A:Reference number: A61093; MUID:86279063
A:Accession: A61093
A:Molecule type: protein
A:Residues: 1-20 <WAI>
A:Note: the amino terminal residue was tentatively identified as Ser
A:Note: 12-Ser was also found
C:Comment: This glue protein is a component of the adhesive plaque of the byssus, the
C:Keywords: hydroxyproline; tandem repeat
F:4,14/Modified site: 3-hydroxyproline (Pro) #status experimental
F:5,9,15,19/Modified site: 4-hydroxyproline (Pro) #status experimental
F:7,17/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental
F:13/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental

Query Match 20.7%; Score 28; DB 2; Length 20;
Best Local Similarity 44.4%; Pred. No. 1.1e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 4 TYKSTVSHP 12
| | | | |
Db 6 TYKPKITYP 14

RESULT 7
PS0251
15K protein 5106 - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
C:Accession: PS0251
R:Tsugita, A.; Kamo, M.
submitted to JPIID, April 1993
A:Reference number: PS0209
A:Accession: PS0251
A:Molecule type: protein
A:Residues: 1-15 <TSU>
A:Experimental source: germ, strain Nihonbare
C:Comment: molecular weight 15K, pI 9.2.

Query Match 20.4%; Score 27.5; DB 2; Length 15;
Best Local Similarity 42.9%; Pred. No. 9.6e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 3; Gaps 1;

QY 9 VSHP---DLPREV 19
| | | | |
Db 2 VTYPINADXPRAI 15

RESULT 8
S78765
ribosomal protein MRP-S24, mitochondrial - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

us-09-701-623c-84.closed.rpr

Mon Aug 26 11:09:50 2002

F:25/Modified site: amidated carboxyl end (Gln) #status experimental

C:Accession: S78765

R:Grack, H.R.

submitted to the Protein Sequence Database, July 1999

A:Reference number: S78760

A:Accession: S78765

A:Molecule type: protein

A:Residues: 1-11 <GRA>

C:Keywords: mitochondrion

F:1-11/Product: ribosomal protein MRP-S24 (fragment) #status experimental <MAT>

Query Match 20.0%; Score 27; DB 2; Length 11;
 Best Local Similarity 30.0%; Pred. No. 8.2e+02;
 Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPDLPREVVR 20

DB 2 HVDVPRDLTK 11

RESULT 9

S41390

p7 protein - human adenovirus 3 (fragment)

C:Species: Mastadenovirus h3 (human adenovirus 3)

C:Date: 06-Jan-1995 #sequence_revision 30-Jan-1998 #text_change 08-Oct-1999

C:Accession: S41390

R:Cuzange, A.; Chroboczek, J.; Jacrot, B.

submitted to the EMBL Data Library, January 1994

A:Description: The penton base of human adenovirus type 3 has the RGD motif.

A:Reference number: S41388

A:Accession: S41390

A:Molecule type: DNA

A:Residues: 1-23 <CUZ>

A:Cross-references: EMBL:Z29487; NID:G444048; PIDN:CAA82623.1; PID:G444051

A:Experimental source: serotype 3

Query Match 20.0%; Score 27; DB 2; Length 23;
 Best Local Similarity 41.7%; Pred. No. 1.8e+03;
 Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHP 12

DB 6 CWEPASSTQHP 17

RESULT 10

A58647

alphaA-conotoxin PIVA [validated] - cone shell (Conus purpurascens)

C:Species: Conus purpurascens (purple cone)

C:Date: 31-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 15-Sep-2000

C:Accession: A58647

R:Hopkins, C.; Grille, M.; Miller, C.; Shon, K.J.; Cruz, L.J.; Gray, W.R.; Dykert, J.;

J. Biol. Chem. 270, 22361-22367, 1995

A:Title: A new family of Conus peptides targeted to the nicotinic acetylcholine receptor

A:Reference number: A58647; MUID:95403432

A:Accession: A58647

A:Molecule type: protein

A:Residues: 1-25 <HOP>

R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.;

submitted to the Brookhaven Protein Data Bank, December 1996

A:Reference number: A67666; PDB:1PIP

A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue

B:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.;

Biochemistry 36, 1669-1677, 1997

A:Title: NMR structure determination of a novel conotoxin, [Pro 7,13] alpha A-conotoxin

A:Reference number: A58646; MUID:97200721

A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR

C:Superfamily: unassigned conotoxins

C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline; pos

F:2-16,3-11,14-23/Disulfide bonds: #status experimental

F:7,13/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F:20/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 19.6%; Score 26.5; DB 2; Length 25;
 Best Local Similarity 41.7%; Pred. No. 2.3e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 CGETYKSTVSHP 12

DB 3 CG-SYPNACHP 13

RESULT 11

I46016

cyokeratin 4 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999

C:Accession: I46016

R:Blessing, M.; Jorcano, J.L.; Franke, W.W.

EMBO J. 8, 117-126, 1989

A:Title: Enhancer elements directing cell-type-specific expression of cyokeratin gen

A:Reference number: I46016; MUID:89231609

A:Accession: I46016

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-9 <BLE>

A:Cross-references: EMBL:X14478; NID:G303; PIDN:CAA32640.1; PID:G577897

Query Match 19.3%; Score 26; DB 2; Length 9;
 Best Local Similarity 83.3%; Pred. No. 2.8e+05;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TYKSTV 9

DB 2 SYKSTV 7

RESULT 12

I38335

hypothetical TEL/MNI mutant fusion protein type II - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000

C:Accession: I38335

R:Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A

Oncogene 10, 1511-1519, 1995

A:Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in f

A:Reference number: I38031; MUID:95249265

A:Accession: I38335

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-15 <BU1>

A:Cross-references: EMBL:X85024; NID:G971471; PIDN:CAA59397.1; PID:G971472

C:Comment: This sequence is the chimeric product of a translocation mutation.

C:Genetics:

A:Gene: ETV6/MNI; TEL/MNI

A:Map position: 22q11/12p13

C:Keywords: fusion protein

Query Match 19.3%; Score 26; DB 4; Length 15;
 Best Local Similarity 54.5%; Pred. No. 1.6e+03;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 YKSTVSHPDLP 15

DB 2 YRSPHSAHDLP 12

RESULT 13

A49255

T-cell receptor beta chain V-D-J-C region (V beta 7, J beta 1.6) - human (fragment)

C:Species: Homo sapiens (man)

C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: A49255
R;Rosenberg, W.M.; Moss, P.A.; Bell, J.I.
Eur. J. Immunol. 22, 541-549, 1992

A;Title: Variation in human T cell receptor V beta and J beta repertoire: analysis using
A;Reference number: A49039; MUID:92164737
C;Accession: A49255
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-16 <ROS>
A;Note: sequence extracted from NCBI backbone (NCBIP:90722)
C;Keywords: T-cell receptor

Query Match 19.3%; Score 26; DB 2; Length 16;
Best Local Similarity 33.3%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGEYKSTVSHP 12
| : | | : |
DB 1 CASSYPGTQNSP 12

RESULT 14
PA0012
superoxide dismutase (EC 1.15.1.1) (Mn) - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 03-May-1996
C;Accession: PA0012
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPD, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensional
A;Reference number: PA0001
A;Accession: PA0012
A;Molecule type: protein
A;Residues: 1-19 <RAM>
A;Experimental source: callus
C;Superfamily: superoxide dismutase (Mn)
C;Keywords: manganese; metalloprotein; oxidoreductase

Query Match 19.3%; Score 26; DB 2; Length 19;
Best Local Similarity 62.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 8 TVSHPDLP 15
| : | | | |
DB 3 TFTLPDLP 10

RESULT 15
A41263
kinase-related transforming protein (hck) (EC 2.7.1.-) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 21-Mar-1996
C;Accession: A41263
R;Lock, P.; Ralph, S.; Stanley, E.; Boulet, I.; Ramsay, R.; Dunn, A.R.
Mol. Cell. Biol. 11, 4363-4370, 1991
A;Title: Two isoforms of murine hck, generated by utilization of alternative translation
A;Reference number: A41263; MUID:91342636
A;Accession: A41263
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-23 <LOC>
C;Keywords: phosphotransferase

Query Match 19.3%; Score 26; DB 2; Length 23;
Best Local Similarity 31.2%; Pred. No. 2.5e+03;
Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 6 KSTVSHPDLPREVRS 21
| : | | | |

Db 4 RSSCEDPGCPRDERA 19
Search completed: August 26, 2002, 10:20:42
Job time: 348 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:30:18 ; Search time 22.95 Seconds
(without alignments)
42.178 Million cell updates/sec

Title: US-09-701-623C-84
Perfect score: 135
Sequence: 1 CGEYKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 1503

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	32	23.7	20	1 CS21_STRTR	P81621 streptococc
2	30	22.2	16	1 YMOR_PSEPU	Q02210 pseudomonas
3	27	20.0	20	1 SODM_HORVU	P28524 condus vul
4	26.5	19.6	25	1 CXAA_CONPU	P55963 condus purpu
5	26	19.3	21	1 CFPA_TREPH	P56738 treponema p
6	25	18.5	10	1 BPP8_BOTIN	P30426 bothrops in
7	25	18.5	13	1 SODM_CANFA	P54712 canis famil
8	25	18.5	14	1 SODM_STRGR	P80732 streptomyc
9	25	18.5	17	1 EFG_THEAQ	Q01697 thermus aqu
10	25	18.5	20	1 UN05_PINPS	P81674 pinus pinas
11	24.5	18.1	20	1 TLI4_SPIOL	P82682 spinacia ol
12	24	17.8	10	1 BPP2_BOTIN	P30422 bothrops in
13	24	17.8	16	1 AU24_LITAU	P82391 litoria aur
14	24	17.8	20	1 VMO2_CHICK	Q9ps49 gallus gall
15	24	17.8	23	1 SODM_RANCA	P36215 rana catesb
16	23	17.0	14	1 CXAL_CONMA	P01521 condus magus
17	23	17.0	15	1 MCA2_RHOOP	P56870 rhodococcus
18	22	17.0	24	1 LEC_CROJU	P16352 crotalaria
19	22	16.3	10	1 ODP2_BOVIN	P11180 bos taurus
20	22	16.3	13	1 BLAC_STRGR	P81173 streptomyc
21	22	16.3	13	1 EP65_HUMAN	P54963 homo sapien
22	22	16.3	15	1 VORA_METTM	P80907 methanobact
23	22	16.3	16	1 ODO2_BOVIN	P11179 bos taurus
24	22	16.3	19	1 UP22_UPEIN	P82028 uperoleia i
25	22	16.3	19	1 UP23_UPEIN	P82029 uperoleia i
26	22	16.3	20	1 HELT_HELHO	P46693 heloderma h
27	22	16.3	20	1 MI17_BOVIN	P35451 bos taurus
28	22	16.3	23	1 GYRA_STRSH	P50071 streptomyc
29	22	16.3	23	1 UHA4_HUMAN	P49289 homo sapien
30	22	16.3	25	1 ANDT_ANDAU	P56684 androctonus
31	21	15.6	11	1 TKN1_PSEGU	P42986 pseudophryn
32	21	15.6	13	1 E121_LITRU	P82097 litoria rub
33	21	15.6	13	1 E122_LITRU	P82098 litoria rub

34	21	15.6	14	1 CXAL_CONCN	P56973 conus conso
35	21	15.6	16	1 LPHI_ECOLI	P03058 escherichia
36	21	15.6	16	1 MMPX_SOLTU	P80501 solanum tub
37	21	15.6	17	1 UP34_UPEMJ	P82041 uperoleia m
38	21	15.6	17	1 UP37_UPEMJ	P82044 uperoleia m
39	21	15.6	19	1 CXA2_CONST	P28879 conus stria
40	21	15.6	20	1 COXM_THUOB	P80981 thunnus obe
41	21	15.6	20	1 COXN_THUOB	P80980 thunnus obe
42	21	15.6	22	1 CYTB_THETS	P81064 theromyzon
43	21	15.6	23	1 XYC1_ACIGB	P46365 acinetobact
44	21	15.6	25	1 FRHB_METBA	P80490 methanosarc
45	20	14.8	13	1 CXAL_CONST	P15471 conus stria

ALIGNMENTS

RESULT 1
ID CS21_STRTR STANDARD; PRT; 20 AA.
AC P81621;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE 21 kDa cold-shock induced protein (Fragment).
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE.
RC STRAIN=PB18;
RA MEDLINE=99456673; PubMed=10525839;
RX Perrin C., Guimont C., Bracqurt P., Gaillard J.L.;
RT "Expression of a new cold shock protein of 21.5 kDa and of the major
cold shock protein by Streptococcus thermophilus after cold shock.";
RL Curr. Microbiol. 39:342-347(1999).
CC -!- INDUCTION: BY COLD SHOCK.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2390 MW; 9C2CA57F266B80AD CRC64;

Query Match 23.7%; Score 32; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 EYKSTVSH 11
|||
DB 6 ETIKETVNH 14

RESULT 2
ID YMOR_PSEPU STANDARD; PRT; 16 AA.
AC Q02210;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in MORA 3'region (Fragment).
OS Pseudomonas putida.
OG plasmid pMDH7.2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M10;
RX MEDLINE=93199531; PubMed=8452544;
RA Willey D.L., Caswell D.A., Lowe C.R., Bruce N.C.;
RT "Nucleotide sequence and over-expression of morphine dehydrogenase, a
plasmid-encoded gene from Pseudomonas putida M10.";
RL Biochem. J. 290:539-544(1993).
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; M94775; AAB17357.1; -

DR PIR; S30384; S30384.

KW Hypothetical protein; Plasmid.

FT NON_TER 16 16

SQ SEQUENCE 16 AA; 1962 MW; A2F1EB8C172766ED CRC64;

Query Match 22.2%; Score 30; DB 1; Length 16;

Best Local Similarity 53.8%; Pred. No. 1.2e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 13 DLPREVVRSTAKC 25

|||| ||: |

Db 5 DLPKRS--RSLRSC 15

RESULT 3

ID SODM_HORVU STANDARD; PRT; 20 AA.

AC P28524;

DT 01-DEC-1992 (Rel. 24, Created)

DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Superoxide dismutase [Mn], mitochondrial (EC 1.15.1.1) (Fragment).

GN SODA.

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;

OC Triticeae; Hordeum.

OX NCBI_TaxID=4513;

RN [1]

RP SEQUENCE.

RC STRAIN=CV. CM 72; TISSUE=Root;

RA Harkman W.J., Tao H.P., Tanaka C.K.;

RT 'Germin'-like polypeptides increase in barley roots during salt

RT stress.";

RL Plant Physiol. 97:366-374(1991).

CC -|- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.

CC -|- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).

CC -|- COFACTOR: Manganese (by similarity).

CC -|- SUBUNIT: HOMOTETRAMER.

CC -|- SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -|- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE

CC FAMILY.

DR HSSP; P04179; 1AP6.

DR Mendel; 16446; HORVU:SodA; mnl6446.

DR InterPro; IPR001189; SOD_MI.

DR Pfam; PF00081; sodfe; 1.

DR PROSITE; PS00088; SOD.MN; PARTIAL.

KW Oxidoreductase; Manganese; Mitochondrion.

FT NON_TER 20 20

SQ SEQUENCE 20 AA; 2152 MW; 7F7CAE8DF1C9657 CRC64;

Query Match

Best Local Similarity 20.0%; Score 27; DB 1; Length 20;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 7 STVSHPDLP 15

:|: ||||

Db ATFTLPDLP 10

RESULT 4

CX44_CONPU
ID CX44_CONPU STANDARD; PRT; 25 AA.
AC P55963;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-A conotoxin PIVA.
OS Conus purpurascens (Purple cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=41690;
RN [1]
RP SEQUENCE.

RC TISSUE=Venom;

RX MEDLINE=95403432; PubMed=7673220;

RA Hopkins C., Griller M., Miller C., Shon K.-J., Cruz L.J., Gray W.R.,

RA Dykert J., Rivier J., Yoshikami D., Olivera B.M.;

RT "A new family of Conus peptides targeted to the nicotinic

RT acetylcholine receptor.";

RL J. Biol. Chem. 270:22361-22367(1995).

RN [2]

RP STRUCTURE BY NMR.

RX MEDLINE=97200721; PubMed=9048550;

RA Han K.-H., Hwang K.-J., Kim S.-M., Kim S.-K., Gray W.R., Olivera B.M.,

RA Rivier J., Shon K.-J.;

RT "NMR structure determination of a novel conotoxin, [Pro 7,13] alpha

RT A-conotoxin PIVA.";

RL Biochemistry 36:1669-1677(1997).

CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY

CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS

CC INHIBIT THEM.

CC -|- SUBCELLULAR LOCATION: Secreted.

CC -|- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.

DR PDB; 1PIP; 07-JUL-97.

KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;

KW Hydroxylation; Venom; 3D-structure.

FT DISULFID 2 16

FT DISULFID 3 11

FT DISULFID 14 23

FT MOD_RES 7 7

FT MOD_RES 13 13

FT MOD_RES 20 20

FT MOD_RES 25 25

SQ SEQUENCE 25 AA; 2608 MW; 9E2147898D697640 CRC64;

Query Match

Best Local Similarity 19.6%; Score 26.5; DB 1; Length 25;

Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 CGETYKSTVSHP 12

||: |:

Db 3 CG-SYPNACHP 13

RESULT 5

ID CFPA_TREPH STANDARD; PRT; 21 AA.

AC P56738;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE Cytoplasmic filament protein A (Fragment).

GN CFPA.

OS Treponema phagedenis.

OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.

OX NCBI_TaxID=162;

RN [1]

RP SEQUENCE.

RC STRAIN=KAZAN 5;

RX MEDLINE=96236033; PubMed=8655496;

RA You Y., Elmore S., Colton L.L., Mackenzie C., Stoops J.K.,

RA Weinstein G.M., Norris S.J.;

RT "Characterization of the cytoplasmic filament protein gene (cfpA) of
 RT *Treponema pallidum* subsp. *pallidum*.";
 RL J. Bacteriol. 178:3177-3187(1996).
 CC -1- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
 CC LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
 CC -1- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
 CC APPosition TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
 CC UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
 KW Structural protein; Antigen.
 FT NON_TER 21
 SQ SEQUENCE 21 AA; 2231 MW; 574604B4FFC2D017 CRC64;

Query Match 19.3%; Score 26; DB 1; Length 21;
 Best Local Similarity 50.0%; Pred. No. 6.8e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 VSHPDLPREV 18
 | | | | |
 Db 11 VFHPKPSAV 20

RESULT 6
 BPP8_BOTIN STANDARD; PRT; 10 AA.
 AC P30426;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Bradykinin-potentiating peptide S5,1 (Angiotensin-converting
 DE enzyme inhibitor).
 OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8723;
 RN [1]
 RP SEQUENCE.

TISSUE=Venom;
 RC MEDLINE=90351557; PubMed=2386615;
 RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
 RT "Primary structure and biological activity of bradykinin potentiating
 RT peptides from Bothrops insularis snake venom.";
 RL J. Protein Chem. 9:221-227(1990).
 CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
 DR PIR; H37196; H37196.
 KW Hypotensive agent; Venom.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 SQ SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 4.2e+02;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPDLP 15
 | | | | |
 Db 5 HPNIP 9

RESULT 7
 SODM_CANFA STANDARD; PRT; 13 AA.
 AC P54712;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Superoxide dismutase [Mn], mitochondrial (EC 1.15.1.1) (Fragment).
 GN SOD2.
 OS *Canis familiaris* (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart;
 RX MEDLINE=98163340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
 RT dog heart proteins.";
 RL Electrophoresis 18:2795-2802(1997).
 CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
 CC FAMILY.
 DR HSP; P04179; IAP6.
 DR HSC-2DPAGE; P54712; DOG.
 DR InterPro; IPR001189; SOD_MI.
 DR Pfam; PF00081; sodfe; 1.
 DR PROSITE; PS00088; SOD_MN; PARTIAL.
 KW Oxidoreductase; Manganese; Mitochondrion.
 FT NON_TER 13
 SQ SEQUENCE 13 AA; 1476 MW; 9C9651DE8BE0672A CRC64;

Query Match 18.5%; Score 25; DB 1; Length 13;
 Best Local Similarity 83.3%; Pred. No. 5.6e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 SHPDLP 15
 | | | | |
 Db 3 SLPDLP 8

RESULT 8
 SODN_STRGR STANDARD; PRT; 14 AA.
 AC P80732;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Superoxide dismutase [Ni] (EC 1.15.1.1) (NISOD) (Nickel-containing
 DE superoxide dismutase) (Fragment).
 GN SODN OR SOD1.
 OS Streptomyces griseus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1911;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=KCTC 9006;
 RX MEDLINE=97056064; PubMed=8900409;
 RA Youn H.-D., Youn H., Lee J.-W., Yim Y.-I., Lee J.K., Hah Y.C.,
 RA Kang S.-O.;
 RT "Unique isozymes of superoxide dismutase in *Streptomyces griseus*.";
 RL Arch. Biochem. Biophys. 334:341-348(1996).
 CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
 CC -1- COFACTOR: NICKEL.
 CC -1- SUBUNIT: HOMOTETRAMER (POTENTIAL).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 KW Oxidoreductase; Nickel.
 FT NON_TER 14
 SQ SEQUENCE 14 AA; 1457 MW; 305C93EF783F2AC8 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 14;
 Best Local Similarity 62.5%; Pred. No. 6.1e+02;
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPDLPREV 18
| | | | |
Db 1 HSDLPSCV 8

RESULT 9

EFG_THEAQ STANDARD; PRT; 17 AA.
AC Q01697;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor G (EF-G) (Fragment).
GN FUSA OR FUS.
OS Thermus aquaticus.
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=271;
RN [1]
SEQUENCE FROM N.A.

RP STRAIN=EP 00276;
RX MEDLINE=92362620; PubMed=1499561;
RA Voss R.H., Hartmann R.K., Lippmann C., Alexander C., Jahn O.,
RA Erdmann V.;
RT "Sequence of the tufA gene encoding elongation factor EF-Tu from
Thermus aquaticus and overproduction of the protein in Escherichia
coli.";
RL Eur. J. Biochem. 207:839-846(1992).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION
OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
RIBOSOME.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
CC EMBL; X66322; CA446997.1; -.
DR PIR; S29294; S29294.
DR HSSP; P13551; 1DAR.
DR InterPro; IPR000795; GTP_EFTU.
DR PROSITE; PS00301; EFACOR_GTP; PARTIAL.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NON_TER 1 1
SQ SEQUENCE 17 AA; 2094 MW; EA46E1EF05F86E1D CRC64;

Query Match 18.5%; Score 25; DB 1; Length 17;
Best Local Similarity 21.4%; Pred. No. 7.6e+02;
Matches 3; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 11 HPDLPREVRSIAK 24
: : : : :
Db 2 YQEVPRQIQEKLK 15

RESULT 10

UN05_PINPS STANDARD; PRT; 20 AA.
AC P81674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of needles (N147) (Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;

RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Prigiero J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
PROTEIN IS: 5.6, ITS MW IS: 36 kDa.
FT NON_TER 1 1
FT NON_CONS 11 12 F -> I.
FT VARIANT 13 13 Y -> E.
FT VARIANT 14 14 Y -> E.
FT VARIANT 15 15 R -> K.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2438 MW; 9F4E4678E086C298 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 20;
Best Local Similarity 80.0%; Pred. No. 9.1e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 PDLPR 16
: : : : :
Db 16 PELPR 20

RESULT 11

TL14_SPIOL STANDARD; PRT; 20 AA.
AC P82682;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Thylakoid lumenal 14.7 kDa protein (P14.7) (Fragment).
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllaceae; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RA Kieselbach T., Pettersson U., Bystedt M., Schroeder W.P.;
RL Submitted (JUL-2000) to the SWISS-PROT data bank.
CC -1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
KW Chloroplast; Thylakoid.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2211 MW; 984B7B4C52F5E3E0 CRC64;

Query Match 18.1%; Score 24.5; DB 1; Length 20;
Best Local Similarity 46.2%; Pred. No. 1.1e+03;
Matches 6; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 6 KSTVSHPD-LPRE 17
: : : : :
Db 1 KTGYNKPELLPKE 13

RESULT 12

BPP2_BOTIN STANDARD; PRT; 10 AA.
AC P30422;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Bradykinin-potentiating peptide S4,3,1 (10C) (Angiotensin-converting
enzyme inhibitor).
OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidae;

OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; B37196; B37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1213 MW; 30C53546C761F773 CRC64;

Query Match 17.8%; Score 24; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 6e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPDLP 15
DB 5 HPQIP 9

RESULT 13
AU24_LITAU
ID AU24_LITAU STANDARD; PRT; 16 AA.
AC P82391;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aurein 2.4 [Contains: Aurein 2.4.1]
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=8371;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
australian bell frogs Litoria aurea and Litoria raniformis the
solution structure of aurein 1.2.";
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -1- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
L.INNOCUA, M.LUTEUS, S.AUREUS, S.EPIDERMIDIS AND S.UBERIS.
CC PROBABLY ACTS BY DISTURBING MEMBRANE FUNCTIONS WITH ITS
CC AMPHIPATHIC STRUCTURE. SHOWS ANTICANCER ACTIVITY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT PEPTIDE 1 16 AUREIN 2.4.
FT PEPTIDE 3 16 AUREIN 2.4.1.
FT MOD_RES 16 16 AMIDATION.
SQ SEQUENCE 16 AA; 1630 MW; 1D87980438AAE2F9 CRC64;

Query Match 17.8%; Score 24; DB 1; Length 16;
Best Local Similarity 45.5%; Pred. No. 1e+03;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 13 DLPREVRSIA 23
DB 4 DIVKKVVGTTA 14

RESULT 14
VMO2_CHICK
ID VMO2_CHICK STANDARD; PRT; 20 AA.
AC Q9PS49;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vitelline membrane outer layer protein II (VMO-II) (VMOII) (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Egg yolk;
RX MEDLINE=92392273; PubMed=1520265;
RA Kido S., Morimoto A., Kim F., Doi Y.;
RT "Isolation of a novel protein from the outer layer of the vitelline
membrane.";
RL Biochem. J. 286:17-22(1992).
CC -1- FUNCTION: EXACT FUNCTION NOT KNOWN, COMPONENT OF THE OUTER
MEMBRANE OF THE VITELLINE LAYER OF THE EGG.
CC -1- PTM: ALL CYSTEINE RESIDUES OF THE MATURE PROTEIN ARE INVOLVED IN
DISULFIDE BONDS.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2325 MW; 45FC7989AB7527C7 CRC64;

Query Match 17.8%; Score 24; DB 1; Length 20;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 14 LPREVRSI 22
DB 1 LPRDTSRXV 9

RESULT 15
SODM_RANCA
ID SODM_RANCA STANDARD; PRT; 23 AA.
AC P36215;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Superoxide dismutase [Mn], mitochondrial (EC 1.15.1.1) (Fragment).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=87126854; PubMed=3492965;
RA Abe Y., Okazaki T.;
RT "Purification and properties of the manganese superoxide dismutase
from the liver of bullfrog, Rana catesbeiana.";
RL Arch. Biochem. Biophys. 253:241-248(1987).
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -1- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) -> O(2) + H(2)O(2).
CC -1- COFACTOR: Manganese.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
FAMILY.
DR HSSP; P04179; IAP6.
DR InterPro; IPR001189; SOD_M1.
DR Pfam; PF00081; sodfe; 1.
DR PROSITE; PS00088; SOD_MN; PARTIAL.
KW Oxidoreductase; Manganese; Mitochondrion.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2594 MW; 5D80ED9B0E04F625 CRC64;

Query Match 17.8%; Score 24; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PDLP 15
 ||||
Db 5 PDLP 8

Search completed: August 26, 2002, 10:30:18
Job time: 638 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:29:35 ; Search time 62.22 seconds
(without alignments)
69.509 Million cell updates/sec

Title: US-09-701-623c-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 8561

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	28.1	24	6 Q9TRX6	Q9trx6 bos taurus
2	33	24.4	21	2 Q9X3C4	Q9x3c4 prochloroco
3	32	23.7	17	6 Q9TR98	Q9tr98 canis famil
4	32	23.7	17	13 O42416	O42416 gallus gall
5	31.5	23.3	17	4 Q16310	Q16310 homo sapien
6	30.5	22.6	21	2 Q9X3J8	Q9x3j8 prochloroco
7	30	22.2	11	2 Q47604	Q47604 escherichia
8	30	22.2	16	6 Q9TR09	Q9tr09 bos taurus
9	30	22.2	20	4 Q9UC08	Q9uce8 homo sapien
10	30	22.2	20	6 Q9TQ26	Q9tq26 macaca fasc
11	30	22.2	23	2 P95839	P95839 staphylococ
12	29	21.5	16	2 Q9R4V4	Q9r4v4 pseudomonas
13	29	21.5	24	6 Q9TRX4	Q9trx4 bos taurus
14	29	21.5	25	8 Q9T2Q0	Q9t2q0 solanum tub
15	28	20.7	24	10 Q941I8	Q941i8 oryza sativ
16	27.5	20.4	20	2 O67964	O67964 prochloroco

17	27.5	20.4	20	2 O67966	O67966 prochloroco
18	27.5	20.4	21	2 Q9WVY2	Q9wvy2 prochloroco
19	27.5	20.4	21	2 Q9R2Y1	Q9r2y1 prochloroco
20	27.5	20.4	21	2 Q9X3D5	Q9x3d5 prochloroco
21	27.5	20.4	21	2 Q9X3D8	Q9x3d8 prochloroco
22	27.5	20.4	25	5 Q9BM56	Q9bm56 lineus sp.
23	27	20.0	15	2 O05991	O05991 staphylococ
24	27	20.0	18	2 Q9X3E9	Q9x3e9 prochloroco
25	27	20.0	18	6 Q28069	Q28069 bos taurus
26	27	20.0	18	6 Q95JA2	Q95ja2 sus scrofa
27	27	20.0	19	2 Q9R7I3	Q9r7i3 prochloroco
28	27	20.0	20	11 Q9QUX8	Q9qux8 rattus sp.
29	27	20.0	21	2 Q9WVZ3	Q9wvz3 prochloroco
30	27	20.0	21	12 Q85667	Q85667 reovirus (t
31	27	20.0	23	12 Q65291	Q65291 human adeno
32	26.5	19.6	21	2 Q9X3C2	Q9x3c2 prochloroco
33	26	19.3	9	6 Q28093	Q28093 bos taurus
34	26	19.3	18	4 Q9UQ08	Q9uq08 homo sapien
35	26	19.3	18	11 Q9JUM8	Q9jjm8 rattus norv
36	26	19.3	19	4 Q16271	Q16271 homo sapien
37	26	19.3	20	2 Q9X3M3	Q9x3m3 prochloroco
38	26	19.3	20	5 Q9U8M9	Q9u8m9 scaptomyza
39	26	19.3	20	13 Q9PRS2	Q9prs2 sphenodon.
40	26	19.3	21	2 Q9X3F4	Q9x3f4 prochloroco
41	26	19.3	21	11 Q9QV41	Q9qval rattus sp.
42	26	19.3	22	4 Q13659	Q13659 homo sapien
43	26	19.3	22	8 Q9GIA0	Q9gia0 sargassum s
44	26	19.3	23	4 Q9HI20	Q9hi20 homo sapien
45	26	19.3	23	8 Q9GIB3	Q9gib3 sargassum c

ALIGNMENTS

RESULT 1

Q9TRX6 PRELIMINARY; PRT; 24 AA.
ID Q9TRX6
AC Q9TRX6
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE GLUTAMATE DEHYDROGENASE (EC 1.4.1.3) (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=91308094; PubMed=1854724;
RA Ozturk D.H., Colman R.F.;
RT "Identification of cysteine-319 as the target amino acid of 8-[(4-bromo-2,3-dioxobutyl)thio]adenosine 5'-triphosphate in bovine liver
RT glutamate dehydrogenase.";
RL Biochemistry 30:7126-7134(1991).
SQ SEQUENCE 24 AA; 2669 MW; D67A84AE4264E740 CRC64;

Query Match 28.1%; Score 38; DB 6; Length 24;
Best Local Similarity 50.0%; Pred. No. 65;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 ETYKSTVSHPDLP 14
Db 7 DTVASTIGHYDI 18

RESULT 2

Q9X3C4 PRELIMINARY; PRT; 21 AA.
ID Q9X3C4
AC Q9X3C4
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070129; AAD20735.1; -.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2318 MW; 78824B529A2C9262 CRC64;

Query Match 24.4%; Score 33; DB 2; Length 21;
Best Local Similarity 46.7%; Pred. No. 3.3e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 STVSHDPLPREVVR 21
||| | :|
DB 2 STLKPKDLRSRKXS 16

RESULT 3
QSTR98
ID Q9TR98 PRELIMINARY; PRT; 17 AA.

AC Q9TR98;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95007849; Pubmed=7923441;
RA Miller W.J., Parmelee D.C., Benjamin T., Sechi S., Dooley K.L.,
RA Kadlubar F.F.;
RT "Plasma proteins as early biomarkers of exposure to carcinogenic
aromatic amines.";
RL Chem.-Biol. Interact. 93:221-234(1994).
SQ SEQUENCE 17 AA; 2024 MW; 1D39F70F7D23B269 CRC64;

Query Match 23.7%; Score 32; DB 6; Length 17;
Best Local Similarity 55.6%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 EYKSTVSH 11
| | | :|
DB 1 EYKSEIAH 9

RESULT 4
O42416
ID O42416 PRELIMINARY; PRT; 19 AA.

AC O42416;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE THROMBOMUCIN (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=97444372; Pubmed=9298993;
RA McNaghy K.M., Pettersson I., Rossi F., Flamme I., Shevchenko A.,
RA Mann M., Graf T.;
RT "Thrombomucin, a novel cell surface protein that defines thrombocytes
and multipotent hematopoietic progenitors.";
RL J. Cell Biol. 138:1395-1407(1997).
DR EMBL; Y13977; CAA74310.1; -.
FT NON_TER 1 19
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2181 MW; 483C3DF97E13EC19 CRC64;

Query Match 23.7%; Score 32; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPDLP 15
|||||
DB 3 HPDLP 7

RESULT 5
Q16310
ID Q16310 PRELIMINARY; PRT; 17 AA.

AC Q16310;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE GC*2 PROTEIN (FRAGMENT).
GN GC*2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95242701; Pubmed=7725672;
RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;
RT "Characterization of mutants of the vitamin-D-binding protein/group
specific component: GC aborigine (1A1) from Australian aborigines and
South African blacks, and 2A9 from south Germany.";
RL Vox Sang. 68:50-54(1995).
DR EMBL; S77130; AAD14250.1; -.
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 23.3%; Score 31.5; DB 4; Length 17;
Best Local Similarity 40.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 12 PD-LPREVVRSAKC 25
|| | :| :|
DB 2 PDATPKELAKLVNKC 16

RESULT 6

Q9X3J8
ID Q9X3J8 PRELIMINARY; PRT; 21 AA.
AC Q9X3J8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;

RT "Genetic diversity in prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream."
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL: AF070190; AAD23228.1; -.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2298 MW; CC946FFCC02C854F CRC64;

Query Match 22.6%; Score 30.5; DB 2; Length 21;
Best Local Similarity 47.4%; Pred. No. 7.8e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 7 STVSHPDLPREWRS-JAK 24
||: ||| :|| :||
Db 2 STLKPKDLADPKLASKLAK 20

RESULT 7

Q47604 Q47604 PRELIMINARY; PRT; 11 AA.
AC Q47604; 1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE REASE PROTEIN (FRAGMENT).
GN REASE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91139577; PubMed=1995588;
RA Tao T., Bourne J.C., Blumenthal R.M.;
RT "A family of regulatory genes associated with type II restriction-
RT modification systems."
J. Bacteriol. 173:1367-1375(1991).
DR EMBL: M63621; AAA24560.1; -.
FT NON_TER 11
SQ SEQUENCE 11 AA; 1296 MW; 3039A71A34472AB7 CRC64;

Query Match 22.2%; Score 30; DB 2; Length 11;
Best Local Similarity 83.3%; Pred. No. 4.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 VSHPD 14
:|||||
Db 1 MSHPD 6

RESULT 8

Q9TR09 Q9TR09 PRELIMINARY; PRT; 16 AA.
AC Q9TR09;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PA700 SUBUNIT (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96216387; PubMed=8621709;
RA DeMartino G.N., Proske R.J., Moomaw C.R., Strong A.A., Song X.,
RA Hisamatsu H., Tanaka K., Slaughter C.A.;
RT "Identification, purification, and characterization of a PA700-
RT dependent activator of the proteasome."
J. Biol. Chem. 271:3112-3118(1996).
SQ SEQUENCE 16 AA; 1888 MW; FF8A93148F3A7BF4 CRC64;

Query Match 22.2%; Score 30; DB 6; Length 16;
Best Local Similarity 60.0%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 11 HPDLPREVVR 20
||||| |
Db 2 HIDLPEQAR 11

RESULT 9

Q9UCE8 Q9UCE8 PRELIMINARY; PRT; 20 AA.
AC Q9UCE8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE D(TTAGGG)N-BINDING PROTEIN B39 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93309464; PubMed=8321232;
RA Ishikawa F., Matunis M.J., Dreyfuss G., Cech T.R.;
RT "Nuclear proteins that bind the pre-mRNA 3' splice site sequence
RT r(UAG/G) and the human telomeric DNA sequence d(TTAGGG)n."
Mol. Cell. Biol. 13:4301-4310(1993).
RL Mol. Cell. Biol. 13:4301-4310(1993).
SQ SEQUENCE 20 AA; 2255 MW; C3CE1955E9A6D210 CRC64;

Query Match 22.2%; Score 30; DB 4; Length 20;
Best Local Similarity 55.6%; Pred. No. 8.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 PDLPREVVR 20
||| | :|
Db 8 PDTPEEKIR 16

RESULT 10

Q9TQ26 Q9TQ26 PRELIMINARY; PRT; 20 AA.
AC Q9TQ26;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=96273610; PubMed=8690030;
RA Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT "Studies on the mechanism of early onset macular degeneration in
RT cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
RT of two proteins in the retina."
Exp. Eye Res. 62:211-219(1996).
RL Exp. Eye Res. 62:211-219(1996).
SQ SEQUENCE 20 AA; 2411 MW; 5F1A6AEB5918F777 CRC64;

Query Match 22.2%; Score 30; DB 6; Length 20;
Best Local Similarity 55.6%; Pred. No. 8.9e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 ETKYSTVSH 11
:|:| | |

Db 11 DTHKSEVAH 19

RESULT 11

P95839 PRELIMINARY; PRT; 23 AA.
 AC P95839;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE ORF56 (FRAGMENT).
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 OX NCBI_TaxID=1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COL;
 RX MEDLINE=96427339; PubMed=8830703;
 RA Wu S., de Lencastre H., Tomasz A.;
 RT "Sigma-B, a putative operon encoding alternate sigma factor of
 RT Staphylococcus aureus RNA polymerase: molecular cloning and DNA
 RT sequencing.";
 RL J. Bacteriol. 178:6036-6042(1996).
 DR EMBL; Y09929; CAA71063.1; -.
 FT NON_TER 1
 SQ SEQUENCE 23 AA; 2541 MW; 7F47717B1767D34F CRC64;

Query Match 22.2%; Score 30; DB 2; Length 23;

Best Local Similarity 60.0%; Pred. No. 1e+03; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGETKSTVS 10

: | | | | |
Db 11 CNETYLSSNS 20

RESULT 12

Q9R4V4 PRELIMINARY; PRT; 16 AA.
 AC Q9R4V4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE AMADORI PRODUCT BINDING PROTEIN (FRAGMENT).
 OS Pseudomonas (fluorescent pseudomonads).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
 OX NCBI_TaxID=286;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95050465; PubMed=7961640;
 RA Gerhardsinger C., Taneda S., Marlon M.S., Monnier V.M.;
 RT "Isolation, purification, and characterization of an Amadori product
 RT binding protein from a Pseudomonas sp. soil strain.";
 RL J. Biol. Chem. 269:27297-27302(1994).
 SQ SEQUENCE 16 AA; 1619 MW; 7330EA1CD1193CD6 CRC64;

Query Match 21.5%; Score 29; DB 2; Length 16;

Best Local Similarity 45.5%; Pred. No. 9.9e+02; Mismatches 2; Indels 0; Gaps 0;

QY 7 STVSHPDLPRE 17

: | | | | |
Db 3 AVVAEPDPAE 13

RESULT 13

Q9TRX4 PRELIMINARY; PRT; 24 AA.
 AC Q9TRX4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE GLUTAMATE DEHYDROGENASE (EC 1.4.1.3) (FRAGMENT).
 OS Bos taurus (Bovine).
 OC Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=91308094; PubMed=1854724;
 RA Ozturk D.H., Colman R.F.;
 RT "Identification of cysteine-319 as the target amino acid of 8-[(4-
 RT bromo-2,3-dioxobutyl)thio]adenosine 5'-triphosphate in bovine liver
 RT glutamate dehydrogenase.";
 RL Biochemistry 30:7126-7134(1991).
 SQ SEQUENCE 24 AA; 2626 MW; DE4ED4B54264E740 CRC64;

Query Match 21.5%; Score 29; DB 6; Length 24;

Best Local Similarity 41.7%; Pred. No. 1.5e+03; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYKSTVSHDLP 14

: | | | | |
Db 7 DTYASTIGXYDI 18

RESULT 14

Q9T2Q0 PRELIMINARY; PRT; 25 AA.
 AC Q9T2Q0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MANGANESE SUPEROXIDE DISMUTASE (FRAGMENT).
 OS Solanum tuberosum (Potato).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95392474; PubMed=7545053;
 RA Fester T., Schuster W.;
 RT "Potato mitochondrial manganese superoxide dismutase is an RNA-binding
 RT protein.";
 RL Biochem. Mol. Biol. Int. 36:67-75(1995).
 CC -|- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
 CC -|- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
 CC -|- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -|- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
 CC FAMILY.
 DR InterPro: IPR001189; SOD_MI.
 DR Pfam: PF00081; sodfe; 1.
 DR ProDom: PD000475; SOD_MI; 1.
 KW Manganese; Oxidoreductase.
 SQ SEQUENCE 25 AA; 2713 MW; 45E0DA2EB3ADB3B3 CRC64;

Query Match 21.5%; Score 29; DB 8; Length 25;

Best Local Similarity 75.0%; Pred. No. 1.6e+03; Mismatches 6; Indels 2; Gaps 0;

QY 8 TVSHDLP 15

: | | | | |
Db 3 TFSLPDLP 10

RESULT 15

Q941I18

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:18:42 ; Search time 79.43 Seconds
(without alignments)
34,960 Million cell updates/sec

Title: US-09-701-623C-84

Perfect score: 135

Sequence: 1 CGETKSTVSHPDLPREVVRSTAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 253524

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	135	100.0	25	AA1980077	Optimised IGE-CH3
2	105	77.8	25	AA1979999	Optimised IGE-CH3
3	94	69.6	25	AA1979998	Optimised IGE-CH3
4	94	69.6	25	AA191212	Modified human IGE
5	94	69.6	25	AA168602	Peptide sequence o
6	89	65.9	25	AA180000	Optimised IGE-CH3
7	74	54.8	25	AA180001	Optimised IGE-CH3
8	63	46.7	20	AAW24102	Canine immunoglobi
9	59	43.7	16	AA150900	Antibody 15A.2 equ
10	58	43.0	22	AA142585	IgE peptide antago
11	58	43.0	24	AA142586	IgE peptide antago

12	55	40.7	17	21	AA150898	Antibody 15A.2 swi
13	54	40.0	17	21	AA150893	Antibody 15A.2 bin
14	54	40.0	17	21	AA150894	Antibody 15A.2 can
15	53	39.3	17	21	AA150897	Antibody 15A.2 fel
16	48	35.6	17	21	AA150896	Antibody 15A.2 gre
17	45	33.3	17	21	AA150895	Antibody 15A.2 hum
18	42	31.1	17	21	AA150894	IgE derived target
19	41	30.4	13	21	AA180048	Human IGE C epsilon
20	41	30.4	13	21	AA180048	Human IGE C epsilon
21	41	30.4	13	21	AA180048	Human IGE C epsilon
22	41	30.4	13	21	AA180048	Human IGE C epsilon
23	39	28.9	22	16	AA172866	Peptide #3 having
24	36	26.7	19	21	AA172867	Peptide #4 having
25	36	26.7	19	21	AA172867	Peptide #4 having
26	36	26.7	19	21	AA172867	Peptide #4 having
27	35	25.9	15	20	AA151033	IgE peptide #11.
28	35	25.9	15	20	AA151033	Rheumatoid arthritis
29	35	25.9	15	20	AA151033	HSV-2 glycoprotein
30	35	25.9	15	20	AA151033	Peptide #10525 enc
31	35	25.9	23	22	AA16830	Human brain expres
32	35	25.9	23	22	AA16830	Human bone marrow
33	35	25.9	23	22	AA16830	Peptide #10886 enc
34	34	25.2	16	22	AA16830	Dog allergen serum
35	34	25.2	16	22	AA16830	Cow milk alpha S1
36	34	25.2	16	22	AA16830	Cow milk alpha S1
37	34	25.2	16	22	AA16830	Cow milk alpha S1
38	34	25.2	16	22	AA16830	Cow milk alpha S1
39	34	25.2	16	22	AA16830	Cow milk alpha S1
40	34	25.2	16	22	AA16830	Cow milk alpha S1
41	34	25.2	16	22	AA16830	Aggrecan degrading
42	34	25.2	16	22	AA16830	Human colon cancer
43	34	25.2	16	22	AA16830	Cow milk alpha S1
44	34	25.2	16	22	AA16830	Cow milk alpha S1
45	34	25.2	16	22	AA16830	Human immunoglobul

ALIGNMENTS

RESULT 1

AA1980077

ID AA1980077 standard; Peptide; 25 AA.

XX

AC AA1980077;

XX

DT 15-MAY-2000 (first entry)

XX

XX Optimised IGE-CH3 domain antigen peptide for horse IgE.

DE

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX Immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Equus caballus.

OS Synthetic.

XX

PN WO9967293-A1.

XX

PD 29-DEC-1999.

XX

PF 21-JUN-1999; 99WO-US13959.

XX

PR 20-JUN-1998; 99US-0100287.

XX

PA (UNBI-) UNITED BIOMEDICAL INC.

XX

XX Wang CV, Walfield AM;

XX

XX WPI; 2000-160578/14.

XX

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions for immunization against allergy

recombinant production of corresponding peptides or in DNA vaccines. Conjugates of (1) that include a promiscuous T helper cell epitope (functional in genetically diverse subjects) in addition to a B cell target epitope, have increased immunogenicity and may include cyclic constraints (disulfide bridge) to stabilize conformational features and maximize cross-reactivity to the natural target. They induce safe (non-anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino acid sequences used in the exemplification of the present invention.

Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 6.1e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy 1 CGETYKSTVSHPDLPREVRSIAKC 25
      |||||:|:|:|:|:|:|:|:|
Db 1 cgetyqsrvtphplpralmrsttkc 25
```

RESULT 4
AAY91212
ID AAY91212 standard; peptide; 25 AA.

DT 22-MAY-2000 (first entry)

DE Modified human IgE CH3 domain, SEQ ID NO:92.

KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin B; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
KW cholesteryl ester transport protein; anti-arteriosclerotic;
KW

OS	Homo sapiens.
OS	Synthetic.

AX
PN
WO9966957-A2.

PD 29-DEC-1999.

21-JUN-1999; 99WO-US13975.

PR 20-JUN-1998; 98US-0100412.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY;

DR WPI; 2000-160564/14.

PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -

PS Example 6; Page 40; 129pp; English.

The invention relates to novel promiscuous T helper cell epitopes (Th), and immunogenic peptides comprising the Th epitopes of the invention along with B cell epitopes. The Th epitopes and peptide immunogens containing them, are used to induce a T helper cell response, specifically against Plasmodium falciparum, cholesterol ester transport protein (CEPP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone

dependent cancer, prevention of boar taint in meat, and immunocastration); for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AAY91121 represents a promiscuous T helper epitope from the measles virus F (VWF) protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the VWF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-Y91243 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AAY90212 is a modified version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th epitope. AAY91223 is a Plasmidom falciuparum circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent CERP-derived peptides and AAY91232-Y91241 are Immunogens comprising a CERP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory invasive protein epitope from *Yersinia* species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 6.1e-08;
Matches 16; Conservative 4; Mismatches 5; Indels

```
Qy      1 CGETYKSTVSHPDLPREVRSIAKC 25
        |||||:|:|:|:|:|:|:|:|
Db      1 cgetvqsrvtphplpralmrsttkc 25
```

RESULT 5
AAY68602
ID AAY68602 standard; peptide: 25 AA.

DT 05-MAY-2000 (first entry)

DE Peptide sequence of the invention.

xx	Helper T cell epitope; peptide immunogen; LHRH;
kw	lutalinising hormone-releasing hormone; spermatogenesis; ovulation;
kw	oestrus; sexual development; sex hormone; promiscuous T helper epitope;
kw	vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
kw	breast cancer; endometriosis; boar taint; meat quality; immunocastration.
xx	
OS	Unidentified..
XX	
PN	WO9966952-A1.
XX	
PD	29-DEC-1999.

```
XX PF 21-JUN-1999; 99WO-US13960.
XX XX
XX PR 20-JUN-1998; 98US-0100414.
XX XX
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX XX
XX PI Wang CY;
XX XX
XX DR WPI; 2000-160562/14.
XX XX
XX PT New peptide immunogen containing luteinising hormone-releasing hormone
XX PT antigen site and helper T cell epitope, for e.g. contraception and
XX PT treatment of cancer -
XX XX
XX PS Disclosure; Page 92; 102pp; English.
XX XX
XX CC The specification describes peptide immunogens comprising a
XX CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
XX CC hormone-releasing hormone (LHRH). The peptide immunogens cause
XX CC induction of a specific immune response to LHRH which is involved in
XX CC regulation of spermatogenesis, ovulation, oestrus, sexual development
XX CC and secretion of sex hormones. Provision of a promiscuous T helper
XX CC epitope (which is functional in genetically diverse subjects) provides
XX CC optimum immunogenicity to the B cell epitopes of the target antigen and
XX CC thus high antibody titres against the target antigen. The peptide
XX CC immunogens of the invention are used to vaccinate against mammalian LHRH,
XX CC for use as (reversible) contraceptive; control of hormone-dependent
XX CC tumours (cancer of prostate or breast, also endometriosis); to prevent
XX CC boar taint (and improve meat quality) and for immunocastration. The
XX CC present sequence appears in the specification.
XX XX
XX SQ Sequence 25 AA;

Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 6.le-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGETYSKSTVSHPDLPREVVRSTAKC 25
DB 1 CGETYSKSTVSHPDLPREVVRSTAKC 25

RESULT 6
AAV80000
ID AAY80000 standard; Peptide; 25 AA.
XX AC AAY80000;
XX XX
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for rat IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;
XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Rattus sp.
XX OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;

Query Match 65.9%; Score 89; DB 21; Length 25;
Best Local Similarity 60.0%; Pred. No. 3.8e-07;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 CGETYSKSTVSHPDLPREVVRSTAKC 25
DB 1 CGEGYQSRVDPHPKPIVRSITKC 25

RESULT 7
AAV80001
ID AAY80001 standard; Peptide; 25 AA.
XX AC AAY80001;
XX XX
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;
XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Mus sp.
XX OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;

WPI; 2000-160578/14.
New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
for immunization against allergy -
Claim 1; Page 99; 155pp; English.
The present invention describes immunoglobulin E (IgE)-CH3 domain
antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
and anti-asthmatic properties. (I) induces polyclonal antibodies
specific for a target effector site on the epsilon-heavy chain of IgE,
and so preventing triggering and activation of mast cells and basophils
and downregulation of IgE synthesis. Conjugates, or fusion peptides,
containing (I) are used for active immunisation against IgE-mediated
allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
dermatitis. Nucleic acids that encode these compounds are useful for
recombinant production of corresponding peptides or in DNA vaccines.
Conjugates of (I) that include a promiscuous T helper cell epitope
(functional in genetically diverse subjects), in addition to a B cell
target epitope, have increased immunogenicity and may include cyclic
constraints (disulfide bridge) to stabilise conformational features and
maximize cross-reactivity to the natural target. They induce safe
(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
acid sequences used in the exemplification of the present invention.
Sequence 25 AA;
```

PS Claim 1; Page 100; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAV7994 to AAV8004 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

SQ

Query Match 54.8%; Score 74; DB 21; Length 25;
 Best Local Similarity 52.0%; Pred. No. 9.3e-05;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHDPLPREVRSIAKC 25
 II III III III III I
 Db 1 cgygysivdrpdkpivrsitlc 25

RESULT 8
 AAW24102
 ID AAW24102 standard; peptide; 20 AA.
 XX AAW24102;
 XX
 XX 21-NOV-1997 (first entry)
 DT Canine immunoglobulin E peptide 5.
 XX Canine immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
 XX Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
 XX Canis familiaris.
 OS JP09169795-A.
 PN 30-JUN-1997.
 XX 22-DEC-1995; 95JP-0334381.
 PF 22-DEC-1995; 95JP-0334381.
 XX (HITB) HITACHI CHEM CO LTD.
 XX WPI; 1997-389423/36.
 DR N-PSDB; AAT85651.
 XX Canine immunoglobulin E peptide fragment and related DNA - useful
 PT for the preparation of anti-canine immunoglobulin E antibody
 XX Claim 2; Page 9; 12pp; Japanese.
 PS AAW24098-106 are peptide fragments containing at least 5 continuous
 XX amino acids of the partial canine immunoglobulin E (IgE) protein shown
 CC in AAW24097. The peptides are used for the preparation of anti-canine
 CC IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
 CC of canine allergies.

XX Sequence 20 AA;

Query Match 46.7%; Score 63; DB 18; Length 20;
 Best Local Similarity 68.8%; Pred. No. 0.004;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSIAK 24
 I:II III::IIIIII
 Db 4 vthpikpdivrsiak 19

RESULT 9
 AAV50900
 ID AAV50900 standard; peptide; 16 AA.
 XX AAV50900;
 AC AAV50900;
 XX 24-FEB-2000 (first entry)
 DT Antibody 15A.2 equine IgE binding epitope 1.
 DE Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX Synthetic.
 OS EP957111-A2.
 PN 17-NOV-1999.
 PD 09-APR-1999; 99EP-0107035.
 PF 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX (IDEX-) IDEXX LAB INC.
 PA Lawton R, Mermer B, Francoeur G;
 XX WPI; 2000-040833/04.
 DR Binding proteins used for treatment or prophylaxis of canine allergy -
 PT Disclosure; Fig 7; 30pp; English.
 XX This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAV50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.

XX Sequence 16 AA;

SQ

Query Match 43.7%; Score 59; DB 21; Length 16;
 Best Local Similarity 91.7%; Pred. No. 0.013;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 KSTVSHDLPRE 17
 I:IIIIIIIIII
 Db 1 kctvshpdlpre 12

RESULT 10
 AAY42585
 ID AAY42585 standard; peptide; 22 AA.
 XX AAY42585;
 AC AAY42585;
 XX 10-JAN-2000 (first entry)
 DT

```
XX DE IgE peptide antagonist.
XX KW Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
KW receptor-binding; binding determinant sequence; anti-IgE antibody;
KW allergic disease.
XX OS Homo sapiens.
XX PN US5965709-A.
XX PD 12-OCT-1999.
XX PF 21-APR-1994; 94US-0232539.
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX PA (GETH ) GENENTECH INC.
XX PI Jardieu PM, Presta LG;
XX WPI; 1999-579941/49.
XX XX Immunoglobulin E variants as peptide antagonists useful for raising and
PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
PT purification of FcepsilonRI receptor and in the treatment of allergic
PT diseases -
XX XX
PS Disclosure; Column 9; 37pp; English.
XX CC The invention provides Immunoglobulin E (IgE) antagonists comprising one
XX or more of the FcepsilonRI receptor-binding determinant sites of human
CC IgE. The antagonists include IgE variants comprising an immunoglobulin
CC template and binding determinant sequences (bds) Cbds, EFbds and the
CC sequence shown in AAY42581. The Cbds (CD loop binding determinant
CC the EFbds (EF loop binding determinant sequence) are selected from
CC the EFbds (EF loop binding determinant sequence) are selected from
CC sequences shown in AAY42578-Y42580. The variants are useful in raising
CC and screening anti-IgE antibodies, in the isolation and purification of
CC FcepsilonRI receptor and in the treatment and prophylaxis of allergic
CC diseases.
XX SQ Sequence 22 AA;
XX XX
XX Query Match 43.0%; Score 58; DB 20; Length 22;
XX Best Local Similarity 55.6%; Pred. No. 0.028;
XX Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 2 GETYKSTVSHPDLPREV 19
XX ||||: |||| ||| :
XX Db 5 getyqcrvthphipralm 22
XX
XX RESULT 11
XX AAY42586
XX ID AAY42586 standard; peptide; 24 AA.
XX AC AAY42586;
XX XX
XX DT 10-JAN-2000 (first entry)
XX DE IgE peptide antagonist.
XX XX
XX KW Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
XX KW receptor-binding; binding determinant sequence; anti-IgE antibody;
XX KW allergic disease.
XX OS Homo sapiens.
XX PN US5965709-A.
XX XX
XX XX
XX Query Match 43.0%; Score 58; DB 20; Length 24;
XX Best Local Similarity 55.6%; Pred. No. 0.031;
XX Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
XX QY 2 GETYKSTVSHPDLPREV 19
XX ||||: |||| ||| :
XX Db 7 getyqcrvthphipralm 24
XX
XX RESULT 12
XX AAY50898
XX ID AAY50898 standard; peptide; 17 AA.
XX AC AAY50898;
XX XX
XX DT 24-FEB-2000 (first entry)
XX DE Antibody 15A.2 swine IgE binding epitope 1.
XX XX
XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
XX KW epitope; prophylaxis; treatment; mimotope.
XX OS Synthetic.
XX XX
XX PN EP957111-A2.
XX XX
XX PD 17-NOV-1999.
XX XX
XX PF 09-APR-1999; 99EP-0107035.
XX PR 09-APR-1998; 98US-0058331.
XX PR 30-MAR-1999; 99US-0281760.
XX XX
XX PA (IDEX-) IDEXX LAB INC.
XX XX
XX PI Lawton R, Mermer B, Francoeur G;
XX XX
XX WPI; 2000-040833/04.
XX XX
```

```

XX Binding proteins used for treatment or prophylaxis of canine allergy -
PT
XX
PS Disclosure; Fig 7; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 17 AA;

Query Match 40.7%; Score 55; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.061;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSI 22
Db 4 vthpdkpilrsi 17

RESULT 13
AAY50893
ID AAY50893 standard; peptide; 17 AA.
XX
AC AAY50893;
XX
DT 24-FEB-2000 (first entry)
DE Antibody 15A.2 binding peptide 10 from Phdc7c phage display library.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 6; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX

```

```

SQ Sequence 17 AA;

Query Match 40.0%; Score 54; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.088;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSI 22
Db 4 vthpdkpilrsi 17

RESULT 14
AAY50894
ID AAY50894 standard; peptide; 17 AA.
XX
AC AAY50894;
XX
DT 24-FEB-2000 (first entry)
DE Antibody 15A.2 canine IgE binding epitope 1.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 17 AA;

Query Match 40.0%; Score 54; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.088;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSI 22
Db 4 vthpdkpilrsi 17

RESULT 15
AAY50897
ID AAY50897 standard; peptide; 17 AA.
XX

```

AC AAY50897;
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 feline IgE binding epitope 1.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
XX
XX 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
XX (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 17 AA;

Query Match 39.3%; Score 53; DB 21; Length 17;
Best Local Similarity 71.4%; Pred. No. 0.13;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 9 VSHPDLPREVVRST 22
Db |:||||| :||||
4 vthpdlplvirsI 17

Search completed: August 26, 2002, 10:18:43
Job time: 699 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: August 26, 2002, 10:29:34 ; Search time 62.22 Seconds
(without alignments)
69.509 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 8561

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	22.0	17	4 Q16310	Q16310 homo sapien
2	31	22.0	18	6 Q95JA2	Q95JA2 sus scrofa
3	30	21.3	19	15 Q90RH8	Q90RH8 human immun
4	29	20.6	16	8 Q9TZR0	Q9TZR0 solanum tub
5	29	20.6	22	11 Q9QW14	Q9QW14 mus sp. pro
6	29	20.6	23	2 P95839	P95839 staphylococ
7	28	19.9	11	13 Q9QWA2	Q9QWA2 gallus gall
8	27	19.1	15	4 Q9UEM3	Q9UEM3 homo sapien
9	27	19.1	18	4 Q13665	Q13665 homo sapien
10	26.5	18.8	17	6 Q9XSG1	Q9XSG1 bos taurus
11	26	18.4	18	4 Q9BQT0	Q9BQT0 homo sapien
12	26	18.4	21	12 Q85667	Q85667 reovirus (t
13	25.5	18.1	23	2 Q9ZEJ4	Q9ZEJ4 anabaena sp
14	25	17.7	14	3 P90342	P90342 saccharomyc
15	25	17.7	15	13 Q90403	Q90403 discopyge o
16	25	17.7	17	2 Q9L8K0	Q9L8K0 enterococcu

17	25	17.7	17	6 Q9TR98	Q9TR98 canis famil
18	25	17.7	19	4 Q16271	Q16271 homo sapien
19	25	17.7	19	13 Q42416	Q42416 gallus gall
20	25	17.7	19	15 Q90RH4	Q90RH4 human immun
21	25	17.7	20	5 Q9U8N5	Q9U8N5 scaptomyza
22	25	17.7	20	5 Q9U8N2	Q9U8N2 scaptomyza
23	25	17.7	21	4 Q9UCL6	Q9UCL6 homo sapien
24	25	17.7	22	10 Q22501	Q22501 nicotiana t
25	25	17.7	23	5 Q95S95	Q95S95 drosophila
26	25	17.7	24	4 Q16061	Q16061 homo sapien
27	25	17.7	25	6 Q9TTG0	Q9TTG0 ateles belz
28	24	17.0	8	6 Q95M23	Q95M23 sus scrofa
29	24	17.0	13	13 P82386	P82386 litorea ran
30	24	17.0	13	13 P82387	P82387 litorea ran
31	24	17.0	16	4 Q9UK83	Q9UK83 homo sapien
32	24	17.0	17	2 Q34216	Q34216 sphingomona
33	24	17.0	17	4 Q9UCS0	Q9UCS0 homo sapien
34	24	17.0	17	11 Q62547	Q62547 mus spretus
35	24	17.0	17	13 P82394	P82394 litorea ran
36	24	17.0	17	13 P82395	P82395 litorea ran
37	24	17.0	17	13 P82396	P82396 litorea ran
38	24	17.0	18	6 Q9GJW3	Q9GJW3 lagenorhync
39	24	17.0	18	6 Q9GJW2	Q9GJW2 lagenorhync
40	24	17.0	18	6 Q9GJW1	Q9GJW1 lagenorhync
41	24	17.0	19	2 P74875	P74875 salmonella
42	24	17.0	19	15 Q90RF4	Q90RF4 human immun
43	24	17.0	20	2 Q9R4J6	Q9R4J6 pseudomonas
44	24	17.0	20	4 Q9UCR9	Q9UCR9 homo sapien
45	24	17.0	20	5 Q46158	Q46158 lumbricus r

ALIGNMENTS

RESULT 1

Q16310 ID Q16310 PRELIMINARY; PRT; 17 AA.
AC Q16310;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE GC*2 PROTEIN (FRAGMENT).
GN GC*2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95242701; PubMed=7725672;
RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;
RT "Characterization of mutants of the vitamin-D-binding protein/group
RT specific component: GC aborigine (1A1) from Australian aborigines and
RT South African blacks, and 2A9 from south Germany.";
RL Vox Sang. 68:50-54(1995).
DR EMBL; S77130; AADI4250.1; -.
FT NON_TER 17
SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 22.0%; Score 31; DB 4; Length 17;
Best Local Similarity 36.4%; Pred. No. 4.8e+02;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 15 PKDIVRSIAKC 25
||: : : ||
Db 6 PRELAKLVNKC 16

RESULT 2
Q95JA2 ID Q95JA2 PRELIMINARY; PRT; 18 AA.
AC Q95JA2;

```
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE INTEGRIN ALPHA 2 SUBUNIT (FRAGMENT).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CARDIOVASCULAR;
RA Markmann A., Kresse H.;
RT "Regulation of VSMC Differentiation.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF330200; AAL09466.1; -.
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 1944 MW; DFEA4C6E14A5B0EF CRC64;

Query Match 22.0%; Score 31; DB 6; Length 18;
Best Local Similarity 45.5%; Pred. No. 5.1e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 15 PKDIVRSIAKC 25
Db 4 PSDVNSLDLC 14

RESULT 3
ID Q90RH8 PRELIMINARY; PRT; 19 AA.
AC Q90RH8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TAT PROTEIN (FRAGMENT).
GN TAT.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96CG12;
RA Taniguchi Y., Takehisa J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., Obengul, M'Pandi M., M'Pele P., Harada Y.,
RA Hayami M., Ichimura H., Parra J.H.;
RT "Genetic Subtypes of HIV Type 1 Based on the vpu/env Sequences in
RT Republic of Congo-Brazzaville.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF127545; AAK84896.1; -.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 2210 MW; 2A83642B89068236 CRC64;

Query Match 21.3%; Score 30; DB 15; Length 19;
Best Local Similarity 53.8%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12 PHLPKDIVRSIAK 24
Db 7 PHSSKDHQNSIPK 19

RESULT 4
ID Q9T2R0 PRELIMINARY; PRT; 16 AA.
AC Q9T2R0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 16, Last annotation update)
DE CYTOCHROME-C REDUCTASE 14 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
```

```
OS Solanum tuberosum (Potato).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX MEDLINE=94198758; PubMed=7764624;
RA Braun H.P., Kruft V., Schmitz U.K.;
RL Planta 193:99-106(1994).
SQ SEQUENCE 16 AA; 1946 MW; BBC625F8E4A4C8E7 CRC64;

Query Match 20.6%; Score 29; DB 8; Length 16;
Best Local Similarity 50.0%; Pred. No. 9.2e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHLPKDI 18
Db 2 HOYLPEDL 9

RESULT 5
Q9QW14 PRELIMINARY; PRT; 22 AA.
ID Q9QW14;
AC Q9QW14;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
DE PROTEIN TYROSINE KINASE JAK1 (FRAGMENT).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93293238; PubMed=8514334;
RA Harpur A.G., Zimlecki A., Wilks A.F., Falk K., Rotzschke O.,
RA Rammensee H.G.;
RT "A prominent natural H-2 Kd ligand is derived from protein tyrosine
RT kinase JAK1.";
RL Immunol. Lett. 35:235-237(1993).
SQ SEQUENCE 22 AA; 2681 MW; D0110BD1FC3C084B CRC64;

Query Match 20.6%; Score 29; DB 11; Length 22;
Best Local Similarity 25.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 4 TYYSRVTHPHLPKDIV 19
Db 7 SYFPEITHIVIKESV 22

RESULT 6
ID P95839 PRELIMINARY; PRT; 23 AA.
AC P95839;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE ORF56 (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RA MEDLINE=96427339; PubMed=8830703;
RA Wu S., de Lencastre H., Tomasz A.;
RT "Sigma-B, a putative operon encoding alternate sigma factor of
```

RT Staphylococcus aureus RNA polymerase: molecular cloning and DNA
sequencing.*; 178:6036-6042(1996).
RL J. Bacteriol. 178:6036-6042(1996).
DR EMBL; Y09929; CAA71063.1; -.
FT NON_TER 1 1
SQ SEQUENCE 23 AA; 2541 MW; 7F47717B1767D34F CRC64;

Query Match 20.6%; Score 29; DB 2; Length 23;
Best Local Similarity 71.4%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGETYS 7
|||
Db 11 CNETYLS 17

RESULT 7
Q90WA2 PRELIMINARY; PRT; 11 AA.
ID Q90WA2
AC Q90WA2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE EXTRACELLULAR FATTY ACID BINDING PROTEIN (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Wang Q., Li N., Deng X., Li H.;
RT "Single Nucleotide Polymorphism Analysis on Encoding Region of
Extracellular Fatty Acid Binding Protein Genes and Their Associations
With the Fattiness Trait in Chicken."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF402001; AAK94062.1; -.
FT NON_TER 1 1
SQ SEQUENCE 11 AA; 1277 MW; 873417F02B187AAA CRC64;

Query Match 19.9%; Score 28; DB 13; Length 11;
Best Local Similarity 83.3%; Pred. No. 8.7e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYS 7
|||
Db 6 GELYLS 11

RESULT 8
Q9UEM3 PRELIMINARY; PRT; 15 AA.
ID Q9UEM3
AC Q9UEM3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE AXONEMAL DYNEIN, HEAVY CHAIN (FRAGMENT).
GN DNAH3.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Maiti A.K., Mattel M.G., Jorissen M., Volz A., Ziegler A.,
RA Bouvagnet P.;
RT "Chromosomal localization of human dynein heavy chain genes."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ132092; CAA10565.1; -.

FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1646 MW; 4EDFDA937C826170 CRC64;

Query Match 19.1%; Score 27; DB 4; Length 15;
Best Local Similarity 44.4%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 16 KDIVRSIAK 24
||: :||
Db 6 KDLAKALAK 14

RESULT 9
Q13665 PRELIMINARY; PRT; 18 AA.
ID Q13665
AC Q13665
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE CARDIAC MYOSIN BINDING PROTEIN (FRAGMENT).
GN MYBP-C.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96083592; PubMed=7493025;
RA Watkins H., Conner D., Thierfelder L., Jarcho J.A., MacRae C.,
RA McKenna W.J., Maron B.J., Seidman J.G., Seidman C.E.;
RT "Mutations in the cardiac myosin binding protein-C gene on chromosome
11 cause familial hypertrophic cardiomyopathy.";
RL Nat. Genet. 11:434-437(1995).
DR EMBL; S80805; AAB35661.1; -.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 1891 MW; 3EC842CE85A19C97 CRC64;

Query Match 19.1%; Score 27; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYSRVT 10
|||
Db 9 CGGIYVCRAT 18

RESULT 10
Q9XSG1 PRELIMINARY; PRT; 17 AA.
ID Q9XSG1
AC Q9XSG1
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE SMCX (FRAGMENT).
GN SMCX.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Poloumienko A., Blecher S.;
RT "Exon-intron structure of SMCX and SMCY genes in bovine and swine."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135448; AAD34440.1; -.
FT NON_TER 1 1
FT NON_TER 17 17
SQ SEQUENCE 17 AA; 1927 MW; 10351B0D516D16F0 CRC64;

```
Query Match      18.8%; Score 26.5; DB 6; Length 17;
Best Local Similarity 42.9%; Pred. No. 2.4e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 12 PHLPKDIVRSIARC 25
   | : | | : | | |
Db 2 PEIPKGVWR-CPKC 14

RESULT 11
Q9BQTO PRELIMINARY; PRT; 18 AA.
AC Q9BQTO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 2.2 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21195339; PubMed=11297743;
RA Holzmann K., Ambrosch I., Elbling L., Micksche M., Berger W.;
RT "A small upstream open reading frame causes inhibition of human major
RT vault protein expression from a ubiquitous mRNA splice variant.";
RL FEBS Lett. 494:199-104(2001).
DR EMBL; AJ291367; CAC35315.1; -.
KW Hypothetical protein.
SQ SEQUENCE 18 AA; 2179 MW; 5D06F9A3F11CB828 CRC64;

Query Match      18.4%; Score 26; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHLP 15
   | | | |
Db 11 PHLP 14

RESULT 12
Q85667 PRELIMINARY; PRT; 21 AA.
AC Q85667;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE OUTER CAPSID SHELL PROTEIN SIGMA-1 (FRAGMENT).
OS Rotavirus (type 3 / strain Dearling).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10886;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81016754; PubMed=6158163;
RA Li J.K., Keene J.D., Scheible P.P., Joklik W.K.;
RT "Nature of the 3'-terminal sequences of the plus and minus strands of
RT the SI gene of reovirus serotypes 1, 2 and 3.";
RL Virology 105:41-51(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81016752; PubMed=7414954;
RA Li J.K., Scheible P.P., Keene J.D., Joklik W.K.;
RT "The plus strand of reovirus gene S2 is identical with its in vitro
RT transcript.";
RL Virology 105:282-286(1980).
DR EMBL; J02325; AAA47262.1; -.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2398 MW; ADE58797A20D5986 CRC64;
```

```
Query Match      18.4%; Score 26; DB 12; Length 21;
Best Local Similarity 45.5%; Pred. No. 3.5e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 12 PHLPKDIVRSI 22
   | | : | | | |
Db 3 PRLREEVRLI 13

RESULT 13
Q9ZEJ4 PRELIMINARY; PRT; 23 AA.
AC Q9ZEJ4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNAJ2 PROTEIN (FRAGMENT).
GN DNAJ2.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7120;
RA Pohl B.;
RL Thesis (1999), University of Bonn, Botanical Institute.
DR EMBL; AJ132709; CAA10746.1; -.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2493 MW; AC50842868C591ED CRC64;

Query Match      18.1%; Score 25.5; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 4.7e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 4 TYYSRV-THP 12
   | | | | : | |
Db 8 TYYSLLGLHP 17

RESULT 14
P90342 PRELIMINARY; PRT; 14 AA.
AC P90342;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORF YBR090C (FRAGMENT).
GN YBR090C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Feldmann H., Mannhaupt G., Schwarzlose C., Vetter I.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95112788; PubMed=7813418;
RA Feldmann H., Aigle M., Aljinovic G., Andre B., Baclet M.C., Barthe C.,
RA Baur A., Becam A.M., Biteau N., Boles E., Brandt T., Brendel M.,
RA Brueckner M., Bussereau F., Christiansen C., Contreras R., Crouzet M.,
RA Cziepluch C., Demolis N., Delaveau T., Dolignon F., Domdey H.,
RA Duesterhus S., Dubois E., Dujon B., El Bakkouri M., Entian K.D.,
RA Feuermann M., Fiers W., Fobo G.M., Fritz C., Gassenhuber H.,
RA Glandsdorff N., Goffeau A., Grivell L.A., De haan M., Hein C.,
RA Herbert C.J., Hollenberg C.P., Holmstrom K., Jacq C., Jacquet M.,
```

```

RA Jauniaux J.C., Jonniaux J.L., Kalliesoe T., Klesau P., Kirchraht L.,
RA Koetter P., Korol S., Liebl S., Logghe M., Lohan A.J.E., Louis E.J.,
RA Li Z.Y., Maat M.J., Mallet L., Mannheim G., Messenguy F., Miosga T.,
RA Molemans F., Mueller S., Nasr F., Obermaier B., Perea J., Pierard A.,
RA Piravandi E., Pohl F.M., Pohl T.M., Pottier S., Proft M., Purnelle B.,
RA Ramezani Rad M., Rieger M., Rose M., Schaaff-Gerstenschlaeger I.,
RA Scherens B., Schwarzlose C., Skala J., Slonimski P.P., Smits P.H.M.,
RA Souciet J.L., Steensma H.Y., Stucka R., Urrestarazu A.,
RA Van der Aart Q.J., Van Dyck L., Vassarotti A., Vetter I.,
RA Vierendeels F., Visser S., Wagner G., de Wergifosse P., Wolfe K.H.,
RA Zagulski M., Zimmermann F.K., Mewes H.W., Kleine K.;
RT "Complete DNA sequence of yeast chromosome II.";
RL EMBO J. 13:5795-5809(1994).
DR EMBL: Z35957; CAA85041.1; -.
DR SGD: S0000294; YBR090C.
FT NON_TER 1
SQ SEQUENCE 14 AA; 1657 MW; C6B3A4A2E8485212 CRC64;

Query Match 17.7%; Score 25; DB 3; Length 14;
Best Local Similarity 83.3%; Pred. No. 3.3e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 KDIVRS 21
Db :|||||
4 RDIVRS 9

RESULT 15
Q90403
ID Q90403 PRELIMINARY; PRT; 15 AA.
AC Q90403;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
DE AGRIN (FRAGMENT).
OS Discopoge ommata (Electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Hypnosqualea; Pristiogryae; Batoides;
OC Torpediniformes; Narcinoidae; Narcinidae; Discopogidae.
OX NCBI_TaxID=7785;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ELECTRIC LOBE;
RX MEDLINE=95164564; PubMed=7860635;
RA Gesemann M., Denzer A.J., Ruegg M.A.;
RT "Acetylcholine receptor-aggregating activity of agrin isoforms and
RT mapping of the active site.";
RL J. Cell Biol. 128:625-636(1995).
DR EMBL: U16146; AAA64486.1; -.
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 1680 MW; 0B881C342FBALD9A CRC64;

Query Match 17.7%; Score 25; DB 13; Length 15;
Best Local Similarity 50.0%; Pred. No. 3.5e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 VTHPHLPKDI 18
Db |||||
1 VTRSHLANEI 10

```

Search completed: August 26, 2002, 10:29:34
Job time: 649 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:18:42 ; Search time 79.43 Seconds
(without alignments)
34.960 Million cell updates/sec

Title: US-09-701-623c-6
Perfect score: 141
Sequence: 1 CGETYSRYTHPLPKDIVRSTAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 253524

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141	100.0	25	21	AAV79999
2	108	76.6	25	21	AAV79998
3	108	76.6	25	21	AAV91212
4	108	76.6	25	21	AAV68602
5	105	74.5	25	21	AAV80000
6	105	74.5	25	21	AAV80077
7	95	67.4	20	18	AAW24102
8	86	61.0	17	21	AAV50893
9	86	61.0	17	21	AAV50894
10	72	51.1	22	20	AAV42585
11	72	51.1	24	20	AAV42586
					Optimised IgE-CH3
					Optimised IgE-CH3
					Modified human IgE
					Peptide sequence o
					Optimised IgE-CH3
					Optimised IgE-CH3
					Canine immunoglobi
					Antibody 15A.2 bin
					Antibody 15A.2 can
					IgE peptide antago
					IgE peptide antago

12	66	46.8	25	21	AAV80001	Optimised IgE-CH3
13	63	44.7	17	21	AAV50896	Antibody 15A.2 gre
14	62	44.0	17	21	AAV50896	Antibody 15A.2 swi
15	60	42.6	17	21	AAV50895	Antibody 15A.2 hum
16	58	41.1	25	21	AAV80048	IgE derived target
17	52	36.9	17	21	AAV50897	Antibody 15A.2 fel
18	51	36.2	13	21	AAV26518	Human IgE C epsilon
19	51	36.2	13	21	AAV16849	Peptide p13b deriv
20	43.5	30.9	12	21	AAV26517	Human IgE C epsilon
21	43.5	30.9	12	21	AAV16848	Peptide p13a deriv
22	42	29.8	19	21	AAV26503	Human IgE C epsilon
23	42	29.8	19	21	AAV16830	Peptide p7X deriv
24	42	29.8	19	21	AAV51033	IgE peptide #11.
25	41	29.1	20	18	AAW24103	Canine immunoglobi
26	40	28.4	17	17	AAW08011	V3 peptide #8. Hu
27	39	27.7	9	21	AAV69599	Monoclonal antibod
28	39	27.7	9	21	AAV69600	Monoclonal antibod
29	39	27.7	9	21	AAV50891	Antibody 15A.2 bin
30	39	27.7	9	21	AAV50892	Antibody 15A.2 bin
31	39	27.7	14	21	AAV69579	Immunogenic peptid
32	38	27.0	12	21	AAV40058	Anti-hIL12 antibod
33	38	27.0	14	21	AAV26519	Human IgE C epsilon
34	38	27.0	14	21	AAV16850	Peptide p13c deriv
35	38	27.0	23	22	ABG20283	Novel human diago
36	37.5	26.6	16	21	AAV24095	Human apoptosis re
37	37.5	26.6	22	20	AAW97830	Human titin fibron
38	37	26.2	12	21	AAV40060	Anti-hIL12 antibod
39	37	26.2	12	21	AAV26516	Human IgE C epsilon
40	37	26.2	12	21	AAV16847	IgE peptide #16.
41	37	26.2	12	22	AAV51038	Peptide which bind
42	37	26.2	17	11	AAU02179	Peptide used to ge
43	37	26.2	19	22	AAU07353	HJ loop peptide Cs
44	37	26.2	20	20	AAW74259	Human protein frag
45	36	25.5	14	22	AAW00660	

ALIGNMENTS

RESULT 1

AAV79999
ID AAV79999 standard; Peptide; 25 AA.

AC AAV79999;

DT 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Canis sp.

OS Synthetic.

PN WO9967293-A1.

PD 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13959.

PR 20-JUN-1998; 98US-0100287.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX PS Claim 1; Page 99; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 141; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.2e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25

Db 1 cgetysrvthplpkdivrsiakc 25

|||||

RESULT 2

AAY79998

ID AAY79998 standard; Peptide: 25 AA.

XX AC AAY79998;

XX DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX PT WPI; 2000-160578/14.

XX PS New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy -

XX CC Claim 1; Page 21; 155pp; English.

XX PS The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

*.

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC containing (I) are used for active immunisation against IgE-mediated

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;

Best Local Similarity 72.0%; Pred. No. 1.3e-09;

Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25

Db 1 cgetysrvthplpkdivrsiakc 25

RESULT 3

AAY91212

ID AAY91212 standard; peptide; 25 AA.

XX AC AAY91212;

XX DT 22-MAY-2000 (first entry)

XX DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;

XX KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

XX KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

XX KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

XX KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

XX KW plasmodium falciparum; circumsporozoite; antimalarial; CFP;

XX KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9966957-A2.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13975.

XX PR 20-JUN-1998; 98US-0100412.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX PT WPI; 2000-160564/14.

XX PS New artificial T helper cell epitope and derived immunogens with target

XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis

XX PT or human immune deficiency virus -

XX PS Example 6; Page 40; 129pp; English.

XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),

XX CC and immunogenic peptides comprising the Th epitopes of the invention

XX CC along with B cell epitopes. The Th epitopes and peptide immunogens

XX CC containing them, are used to induce a T helper cell response,

XX CC specifically against Plasmodium falciparum, cholesteryl ester transport

CC protein (CTPP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4-CDR2-like domain antigenic site, and AAY91209-Y90211 are a
CC epitope/CD4-CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC epitope, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CPTP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CPTP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.

XX Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 1.3e-09;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| ||||| :||| ||
Db 1 cgetysrvthphlpralmrsttkc 25

RESULT 4
AAY68602
ID AAY68602 standard; peptide; 25 AA.

XX

AC AAY68602;

XX 05-MAY-2000 (first entry)

DE Peptide sequence of the invention.

XX

KW Helper T cell epitope; peptide immunogen; LHRH;

KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;

KW vaccine; sexual development; sex hormone; promiscuous T helper epitope;

KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;

KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX

OS Unidentified.

XX WO9966952-A1.

PN

XX 29-DEC-1999.

PD

XX

PF 21-JUN-1999; 99WO-US13960.

XX

PR 20-JUN-1998; 98US-0100414.

XX

XX (UNBI-) UNITED BIOMEDICAL INC.

PA

XX Wang CY;

PI

XX WPI; 2000-160562/14.

XX

XX New peptide immunogen containing luteinising hormone-releasing hormone

PT antigen site and helper T cell epitope, for e.g. contraception and

PT treatment of cancer

XX

XX Disclosure; Page 92; 102pp; English.

XX

XX The specification describes peptide immunogens comprising a

CC synthetic helper T cell (Th) epitope and a target antigen, luteinising

CC hormone-releasing hormone (LHRH). The peptide immunogens cause

CC induction of a specific immune response to LHRH which is involved in

CC regulation of spermatogenesis, ovulation, oestrus, sexual development

CC and secretion of sex hormones. Provision of a promiscuous T helper

CC epitope (which is functional in genetically diverse subjects) provides

CC optimum immunogenicity to the B cell epitopes of the target antigen and

CC thus high antibody titres against the target antigen. The peptide

CC immunogens of the invention are used to vaccinate against mammalian LHRH,

CC for use as (reversible) contraceptive; control of hormone-dependent

CC tumours (cancer of prostate or breast, also endometriosis); to prevent

CC boar taint (and improve meat quality) and for immunocastration. The

CC present sequence appears in the specification.

XX

SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 1.3e-09;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| ||||| :||| ||
Db 1 cgetysrvthphlpralmrsttkc 25

RESULT 5
AAY80000
ID AAY80000 standard; Peptide; 25 AA.

XX

AC AAY80000;

XX

XX 15-MAY-2000 (first entry)

XX

XX Optimised IgE-CH3 domain antigen peptide for rat IgE.

XX

KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX

OS Rattus sp.

OS Synthetic.

XX

XX WO9967293-A1.

PN

XX 29-DEC-1999.

PD

XX

PF 21-JUN-1999; 99WO-US13959.

```

XX 20-JUN-1998; 98US-0100287.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY, Walfield AM;
XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Claim 1; Page 99; 155pp; English.
XX
XX The present invention describes Immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX
XX Sequence 25 AA;

Query Match 74.5%; Score 105; DB 21; Length 25;
Best Local Similarity 76.0%; Pred. No. 3.8e-09;
Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB 1 CGEGYQSRVDPHPKPLVRSIAKC 25

RESULT 6
AAY80077
ID AAY80077 standard; Peptide; 25 AA.
XX
XX AAY80077;
XX
XX 15-MAY-2000 (first entry)
XX
XX Optimised IgE-CH3 domain antigen peptide for horse IgE.
XX
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX antibody; allergy; allergic disease; immunisation; anti-allergic;
XX anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
XX Equus caballus.
XX Synthetic.
XX
XX WO9967293-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13959.
XX
XX 20-JUN-1998; 98US-0100287.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY, Walfield AM;

```

```

XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Claim 1; Page 146; 155pp; English.
XX
XX The present invention describes Immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX
XX Sequence 25 AA;

Query Match 74.5%; Score 105; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 3.8e-09;
Matches 18; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB 1 CGETYYSRVTHPLPKDIVRSIAKC 25

RESULT 7
AAW24102
ID AAW24102 standard; peptide; 20 AA.
XX
XX AAW24102;
XX
XX 21-NOV-1997 (first entry)
XX
XX Canine immunoglobulin E peptide 5.
XX
XX Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
XX
XX Canis familiaris.
XX
XX JP09169795-A.
XX
XX 30-JUN-1997.
XX
XX 22-DEC-1995; 95JP-0334381.
XX
XX 22-DEC-1995; 95JP-0334381.
XX
XX (HITB ) HITACHI CHEM CO LTD.
XX
XX WPI; 1997-389423/36.
XX N-PSDB; AAT85651.
XX
XX Canine immunoglobulin E peptide fragment and related DNA - useful
XX for the preparation of anti-canine immunoglobulin E antibody
XX
XX Claim 2; Page 9; 12pp; Japanese.
XX
XX AAW24098-106 are peptide fragments containing at least 5 continuous
XX amino acids of the partial canine immunoglobulin E (IgE) protein shown
XX in AAW24097. The peptides are used for the preparation of anti-canine

```

```
CC IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
CC of canine allergies.
XX
SQ Sequence 20 AA;

Query Match 67.4%; Score 95; DB 18; Length 20;
Best Local Similarity 94.7%; Pred. No. 9 9e-08;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 YSRVTHPLPKDIVRSIAK 24
Db 1 ycrvthplpkdivrsiak 19

RESULT 8
AAV50893
ID AAV50893 standard; peptide; 17 AA.
XX
AC AAV50893;
XX
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 binding peptide 10 from PhdC7c phage display library.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
XX (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
XX WPI; 2000-040833/04.
XX
DR Binding proteins used for treatment or prophylaxis of canine allergy -
XX
PT Disclosure; Fig 6; 30pp; English.
XX
PS This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAV50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
XX Sequence 17 AA;

Query Match 61.0%; Score 86; DB 21; Length 17;
Best Local Similarity 94.1%; Pred. No. 2e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6 YSRVTHPLPKDIVRSI 22
Db 1 ycrvthplpkdivrsi 17

RESULT 10
AAV42585
ID AAV42585 standard; peptide; 22 AA.
XX
AC AAV42585;
XX
XX
DT 10-JAN-2000 (first entry)
XX
DE IgE peptide antagonist.
XX
KW Immunoglobulin E; IgE; antagonist; FcεpsilonRI receptor; human; bds;
KW receptor-binding; binding determinant sequence; anti-IgE antibody;
XX allergic disease.
XX
XX Homo sapiens.
XX
XX OS
XX
XX PN US5965709-A.
```

XX PD 12-OCT-1999.
XX PF 21-APR-1994; 94US-0232539.
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX PA (GETH) GENENTECH INC.
XX PI Jardieu PM, Presta LG;
XX DR WPI; 1999-579941/49.
XX PT Immunoglobulin E variants as peptide antagonists useful for raising and
PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
PT purification of FcepsilonRI receptor and in the treatment of allergic
PT diseases -
XX PS Disclosure; Column 9; 37pp; English.
XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one
CC or more of the FcepsilonRI receptor-binding determinant sites of human
CC IgE. The antagonists include IgE variants comprising an immunoglobulin
CC template and binding determinant sequences (bds) CDBds, EFBds and the
CC sequence shown in AAY42581. The CDBds (CD loop binding determinant
CC sequence) are selected from the sequences shown in AAY42567-Y42577 and
CC the EFBds (EF loop binding determinant sequence) are selected from
CC sequences shown in AAY42578-Y42580. The variants are useful in raising
CC and screening anti-IgE antibodies, in the isolation and purification of
CC FcepsilonRI receptor and in the treatment and prophylaxis of allergic
CC diseases.
XX SQ Sequence 22 AA;
Query Match 51.1%; Score 72; DB 20; Length 22;
Best Local Similarity 66.7%; Pred. No. 0.00037;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 GETYYSRVTHPLPKDIV 19
Db |||| |||||: :
5 getyqrvthplpralm 22
RESULT 11
AAV42586
ID AAY42586 standard; peptide; 24 AA.
XX AC AAY42586;
XX DT 10-JAN-2000 (first entry)
XX DE IgE peptide antagonist.
XX KW Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
KW receptor-binding; binding determinant sequence; anti-IgE antibody;
KW allergic disease.
XX OS Homo sapiens.
XX PN US5965709-A.
XX PD 12-OCT-1999.
XX PF 21-APR-1994; 94US-0232539.
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX PA (GETH) GENENTECH INC.
XX PI Jardieu PM, Presta LG;

XX DR WPI; 1999-579941/49.
XX PF Immunoglobulin E variants as peptide antagonists useful for raising and
PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
PT purification of FcepsilonRI receptor and in the treatment of allergic
PT diseases -
XX PS Disclosure; Column 9; 37pp; English.
XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one
CC or more of the FcepsilonRI receptor-binding determinant sites of human
CC IgE. The antagonists include IgE variants comprising an immunoglobulin
CC template and binding determinant sequences (bds) CDBds, EFBds and the
CC sequence shown in AAY42581. The CDBds (CD loop binding determinant
CC sequence) are selected from the sequences shown in AAY42567-Y42577 and
CC the EFBds (EF loop binding determinant sequence) are selected from
CC sequences shown in AAY42578-Y42580. The variants are useful in raising
CC and screening anti-IgE antibodies, in the isolation and purification of
CC FcepsilonRI receptor and in the treatment and prophylaxis of allergic
CC diseases.
XX SQ Sequence 24 AA;
Query Match 51.1%; Score 72; DB 20; Length 24;
Best Local Similarity 66.7%; Pred. No. 0.0004;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 2 GETYYSRVTHPLPKDIV 19
Db |||| |||||: :
7 getyqrvthplpralm 24
RESULT 12
AAV80001
ID AAY80001 standard; Peptide; 25 AA.
XX AC AAY80001;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Mus sp.
XX OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of Immunoglobulin E, fusions
PT for immunization against allergy -
XX PS Claim 1; Page 100; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 46.8%; Score 66; DB 21; Length 25;
 Best Local Similarity 56.0%; Pred. No. 0.0035;
 Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25
 II I I I I I I I I I I I
 Db 1 cgygvgisvdrpdpkpvirsitlc 25

RESULT 13

AAY50896
 ID AAY50896 standard; peptide; 17 AA.

XX AC AAY50896;

XX DT 24-FEB-2000 (first entry)

XX DE Antibody 15A.2 green monkey IgE binding epitope 1.

XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 XX epitope; prophylaxis; treatment; mimotope.

XX OS Synthetic.

XX PN EP957111-A2.

XX PD 17-NOV-1999.

XX PF 09-APR-1999; 99EP-0107035.

XX PR 09-APR-1998; 98US-0058331.

XX PR 30-MAR-1999; 99US-0281760.

XX PA (IDEX-) IDEXX LAB INC.

XX PI Lawton R, Mermer B, Francoeur G;

XX DR WPI; 2000-040833/04.

XX PT Binding proteins used for treatment or prophylaxis of canine allergy -
 XX Disclosure; Fig 7; 30pp; English.

XX This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.

SQ Sequence 17 AA;

Query Match 44.7%; Score 63; DB 21; Length 17;
 Best Local Similarity 78.6%; Pred. No. 0.0065;
 Matches 11; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8 RVTHPHLPKDIVRS 21
 I I I I I I I I I I I I I
 Db 3 rvthphlpralvrs 16

RESULT 14

AAY50898
 ID AAY50898 standard; peptide; 17 AA.

XX AC AAY50898;

XX DT 24-FEB-2000 (first entry)

XX DE Antibody 15A.2 swine IgE binding epitope 1.

XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 XX epitope; prophylaxis; treatment; mimotope.

XX OS Synthetic.

XX PN EP957111-A2.

XX PD 17-NOV-1999.

XX PF 09-APR-1999; 99EP-0107035.

XX PR 09-APR-1998; 98US-0058331.

XX PR 30-MAR-1999; 99US-0281760.

XX PA (IDEX-) IDEXX LAB INC.

XX PI Lawton R, Mermer B, Francoeur G;

XX DR WPI; 2000-040833/04.

XX PT Binding proteins used for treatment or prophylaxis of canine allergy -
 XX Disclosure; Fig 7; 30pp; English.

XX This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.

SQ Sequence 17 AA;

Query Match 44.0%; Score 62; DB 21; Length 17;
 Best Local Similarity 70.6%; Pred. No. 0.0092;
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 YSRVTHPHLPKDIVRSI 22
 I I I I I I I I I I I I I
 Db 1 ycnvthpdlpkpilsrsi 17

RESULT 15

AAY50895
 ID AAY50895 standard; peptide; 17 AA.

XX XX

AC AAY50895;
 XX
 DT 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 human IgE binding epitope 1.
 XX
 KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.
 XX
 PF 09-APR-1999; 99EP-0107035.
 XX
 PR 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Lawton R, Mermer B, Francoeur G;
 XX
 DR WPI; 2000-040833/04.
 XX
 PT Binding proteins used for treatment or prophylaxis of canine allergy -
 PS Disclosure; Fig 7; 30pp; English.
 XX
 CC This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 SQ Sequence 17 AA;

Query Match 42.6%; Score 60; DB 21; Length 17;
 Best Local Similarity 71.4%; Pred. No. 0.019;
 Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 RVTHPHLPKDIVRS 21
 |||||: :||
 Db 3 rvthphlpralmrs 16

Search completed: August 26, 2002, 10:18:42
 Job time: 698 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:19:36 ; Search time 33.13 Seconds
(without alignments)
18.432 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGEYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 134663

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108	76.6	25	3	US-09-100-414B-95
2	108	76.6	25	4	US-09-303-323-95
3	72	51.1	22	2	US-08-232-539D-19
4	72	51.1	24	2	US-08-232-539D-20
5	40	28.4	17	1	US-08-218-025A-186
6	36	25.5	18	2	US-09-017-205-52
7	34.5	24.5	16	1	US-08-077-797A-14
8	34.5	24.5	16	5	PCT-US94-01238-14
9	34	24.1	18	4	US-09-177-249-298
10	32.5	23.0	15	2	US-08-671-094B-8
11	32.5	23.0	17	1	US-08-311-307B-9
12	32	22.7	17	2	US-08-847-696A-9
13	31.5	22.3	24	2	US-08-637-759B-313
14	31.5	22.3	24	3	US-08-871-355A-313
15	31.5	22.3	24	4	US-09-201-945-313
16	31	22.0	24	3	US-08-592-500-27
17	31	22.0	24	3	US-08-195-006-27
18	31	22.0	24	5	PCT-US94-0764A-27
19	30.5	21.6	21	2	US-07-746-705A-11
20	30.5	21.6	21	1	US-08-380-182-10
21	30	21.3	13	1	US-08-466-285-6
22	30	21.3	20	4	US-09-446-504-9
23	30	21.3	20	4	US-09-712-266-9
24	30	21.3	21	2	US-08-480-190-80
25	30	21.3	21	2	US-08-488-379-80
26	30	21.3	21	5	PCT-US93-07545-80
27	30	21.3	23	2	US-08-244-951A-3

28	30	21.3	23	4	US-09-029-348-6	Sequence 6, Appli
29	30	21.3	25	3	US-08-822-324-14	Sequence 14, Appl
30	29	20.6	10	4	US-09-385-442-18	Sequence 18, Appl
31	29	20.6	16	1	US-08-260-582-41	Sequence 41, Appl
32	29	20.6	16	4	US-08-602-999A-226	Sequence 226, Appl
33	29	20.6	16	5	PCT-US95-05471-41	Sequence 41, Appl
34	29	20.6	17	2	US-08-982-597A-24	Sequence 24, Appl
35	29	20.6	17	3	US-09-136-218-24	Sequence 24, Appl
36	29	20.6	18	2	US-09-017-205-9	Sequence 9, Appli
37	29	20.6	18	4	US-09-029-052-5	Sequence 5, Appli
38	29	20.6	18	4	US-08-602-999A-371	Sequence 371, App
39	29	20.6	19	1	US-08-116-733-12	Sequence 12, Appl
40	29	20.6	19	1	US-08-469-615-9	Sequence 9, Appli
41	29	20.6	19	1	US-08-466-763-9	Sequence 9, Appli
42	29	20.6	19	2	US-08-411-142A-9	Sequence 9, Appli
43	29	20.6	24	1	US-07-976-358-24	Sequence 24, Appl
44	29	20.6	25	1	US-07-976-358-15	Sequence 15, Appl
45	29	20.6	25	1	US-07-976-358-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 76.6%; Score 108; DB 3; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGEYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| :||| ||
Db 1 CGEYYSRVTHPLPKDIVRSIAKC 25

```
RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHR PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-95

Query Match 76.6%; Score 108; DB 4; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
    ||||| ||||| :||| ||
Db 1 CGETYQSRVTHPLPRALMRSTTKC 25

RESULT 3
US-08-232-539D-19
; Sequence 19, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

;
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

;
; Query Match 51.1%; Score 72; DB 2; Length 22;
; Best Local Similarity 66.7%; Pred. No. 2.8e-05;
; Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIV 19
    |||| ||||| :|
Db 5 GETYQSRVTHPLPRALM 22

RESULT 4
US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
```


INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-232-539D-20

Query Match 51.1%; Score 72; DB 2; Length 24;
Best Local Similarity 66.7%; Pred. No. 3.1e-05;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPHLPKDIV 19
DB 7 GETYQCRVTHPHLPRAIM 24

RESULT 5

US-08-218-025A-186
Sequence 186, Application US/08218025A
Patent No. 5556744
GENERAL INFORMATION:
APPLICANT: Weiner, David B.
APPLICANT: Ugen, Kenneth E.
APPLICANT: Williams, William V.
TITLE OF INVENTION: Methods and Compositions for Diagnosing
TITLE OF INVENTION: and Treating Certain HIV Infected Patients
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSER: Howson and Howson
STREET: P.O. Box 457, 321 No. 5556744,ristown Road
CITY: Spring House
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19477

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,025A
FILING DATE: 24-MAR-1994

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,451
FILING DATE: 29-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.

REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST33A
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-218-025A-186

Query Match 28.4%; Score 40; DB 1; Length 17;
Best Local Similarity 43.8%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPHLPK 16
DB 1 CGTTFVARGPGIHSK 16

RESULT 6

US-09-017-205-52
Sequence 52, Application US/09017205
Patent No. 5965357
GENERAL INFORMATION:
APPLICANT: Marsden, Howard S.
TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN
TITLE OF INVENTION: DIAGNOSIS OF HERPES SIMPLEX VIRUS TYPE 2
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSER: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 5965357th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/017,205
FILING DATE: 02-FEB-1998
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C.
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 604-436
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100

INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide from HSV-2 glycoprotein G
FRAGMENT TYPE: internal
US-09-017-205-52

Query Match 25.5%; Score 36; DB 2; Length 18;
Best Local Similarity 75.0%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 THPHLPKD 17
DB 7 THPHGPAD 14

RESULT 7

US-08-077-797A-14
Sequence 14, Application US/08077797A
Patent No. 5679548
GENERAL INFORMATION:
APPLICANT: Barbas, Carlos F.
APPLICANT: Rosenblum, Jonathan
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSER: The Scripps Research Institute, Office of
ADDRESSER: Patent Counsel
STREET: 10666 No. 5679548th Torrey Pines Road, TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,797A
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCR1276P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-077-797A-14

Query Match 24.5%; Score 34.5; DB 1; Length 16;
Best Local Similarity 43.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 2 GETYSRVTHPLPKD 17
| : | : | : | : |
Db 1 GDTHRGHLRH-HLPHD 15

RESULT 8
PCT-US94-01238-14
; Sequence 14, Application PC/TUS9401238
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
; TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
; NUMBER OF SEQUENCES: 65
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01238
; FILING DATE: 01-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,797
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
PCT-US94-01238-14

Query Match 24.5%; Score 34.5; DB 5; Length 16;
Best Local Similarity 43.8%; Pred. No. 24;
Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 2 GETYSRVTHPLPKD 17
| : | : | : | : |
Db 1 GDTHRGHLRH-HLPHD 15

RESULT 9
US-09-177-249-298
; Sequence 298, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Ohad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadegari, Ramin
; APPLICANT: Margossian, Linda
; APPLICANT: Harada, John
; APPLICANT: Goldberg, Robert B.
; TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
; TITLE OF INVENTION: Development in Plants
; FILE REFERENCE: 023070-086120US
; CURRENT APPLICATION NUMBER: US/09/177,249
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US 09/071,838
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 298
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-177-249-298

Query Match 24.1%; Score 34; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 16 KDIVRSIAKC 25
| : | : | : | : |
Db 2 KSVVRNVQKC 11

RESULT 10
US-08-671-094B-8
; Sequence 8, Application US/08671094B
; Patent No. 5912232
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Anti-Inflammatory Polypeptide
; TITLE OF INVENTION: Antagonists of Human Interleukin-8
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,094B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63086FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4..6
OTHER INFORMATION: /note= "The cysteine residues
may be substituted with aminobutyric acid, homocysteine or
diaminosuberic acid."
US-08-671-094B-8

Query Match 23.0%; Score 32.5; DB 2; Length 15;
Best Local Similarity 53.8%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 CGETYSRVTHPH 13
| : | | : | | |
Db 4 CIKT-YSKPFHPH 15

RESULT 11
US-08-311-307B-9
; Sequence 9, Application US/08311307B
; Patent No. 5627156
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human
; INTERLEUKIN-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,307B
; FILING DATE: 23-SEP-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues may
; be substituted with aminobutyric acid, homocysteine or
; diaminosuberic acid."

US-08-311-307B-9

Query Match 23.0%; Score 32.5; DB 1; Length 17;
Best Local Similarity 53.8%; Pred. No. 54;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 CGETYSRVTHPH 13
| : | | : | | |
Db 6 CIKT-YSKPFHPH 17

RESULT 12
US-08-847-696A-9
; Sequence 9, Application US/08847696A
; Patent No. 5872726
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist For Human Interleukin-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,696A
; FILING DATE: 28-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/311,307
; FILING DATE: 23-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rigaut, Kathleen D.
; REGISTRATION NUMBER: P 43,047
; REFERENCE/DOCKET NUMBER: 63085C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The Xaa in positions 4 and 6
; in the peptide may be aminobutyric acid, homocysteine, cys
; diaminosuberic acid."

Query Match 22.7%; Score 32; DB 2; Length 17;
Best Local Similarity 62.5%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 YSRVTHPH 13
| : | | : | | |
Db 10 YSKPFHPH 17

RESULT 13
US-08-637-759B-313
; Sequence 313, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 313:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-637-759B-313

Query Match 22.3%; Score 31.5; DB 2; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 4 TYYSRVTHPLPKD 17
|||:|:|:|:
DB 6 TYLSDTN-HLPAE 18

RESULT 14
US-08-871-355A-313
; Sequence 313, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 313:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-871-355A-313

Query Match 22.3%; Score 31.5; DB 3; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

OY 4 TYYSRVTHPLPKD 17
|||:|:|:|:
DB 6 TYLSDTN-HLPAE 18

RESULT 15
US-09-201-945-313
; Sequence 313, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (404) 873-8794
 TELEFAX: (404) 873-8795
 INFORMATION FOR SEQ ID NO: 313:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 24 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 US-09-201-945-313

Query Match 22.3%; Score 31.5; D3 4; Length 24;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 4 TYYSRVTHPLPKD 17
 Db 6 TYLSDITN-HLP AE 18

Search completed: August 26, 2002, 10:19:36
 Job time: 392 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:20:41 ; Search time 43.94 Seconds
(without alignments)
54.671 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPFPKPIVRSITKC 23
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues;
Total number of hits satisfying chosen parameters: 4981

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	22.8	23	2 S48156	alpha-amylase inhi
2	29	20.0	18	2 S26664	microtubule-associ
3	29	20.0	19	2 B53145	high conductance c
4	28	19.3	13	2 A05174	tryptophyllin-13 -
5	28	19.3	23	2 E39855	paralytic peptide
6	28	19.3	23	2 C39855	paralytic peptide
7	28	19.3	23	2 D39855	paralytic peptide
8	27	18.6	13	2 S21152	tryptophyllin-rela
9	27	18.6	24	4 S09363	hypothetical MTCOI
10	26.5	18.3	20	2 S09025	carboxylesterase (
11	26	17.9	18	2 I78841	thrombopoietin rec
12	26	17.9	20	2 A41439	acid ribonuclease
13	26	17.9	22	2 JN0910	N4-(beta-N-acetyl
14	26	17.9	23	2 A47415	mannose-1-phosphat
15	26	17.9	24	2 S29749	serum albumin - do
16	25.5	17.6	25	2 S03456	T-cell receptor al
17	25	17.2	13	2 B47415	mannose-1-phosphat
18	25	17.2	16	2 I51879	cystathionine beta
19	25	17.2	19	2 T02624	hypothetical prote
20	25	17.2	20	2 PNO133	pepsin (EC 3.4.23.
21	25	17.2	20	2 A61093	glue protein - Cal
22	25	17.2	21	2 I54351	gene HEXA protein
23	25	17.2	23	2 S41390	p7 protein - human
24	25	17.2	23	2 S72535	probable acr-2 reg
25	25	17.2	23	2 I40692	cenA protein (IgaI
26	25	17.2	25	2 A60704	serine proteinase
27	24	16.6	14	2 PHI597	Ig H chain V-D-J r
28	24	16.6	16	2 T37075	hypothetical prote
29	24	16.6	18	2 B49254	TCR C gamma 1 chai

30	24	16.6	18	2 B32473	histidine-rich pro
31	24	16.6	18	2 I46653	T-cell receptor de
32	24	16.6	20	2 S58382	hypothetical prote
33	24	16.6	21	2 PQ0257	microbial serine p
34	24	16.6	21	2 I50535	calmodulin - elect
35	24	16.6	22	2 I50533	calmodulin - elect
36	24	16.6	22	2 S73389	hypothetical prote
37	24	16.6	23	2 A60423	monophenol monooxy
38	24	16.6	23	2 A32473	histidine-rich pro
39	24	16.6	25	2 A18864	enkephalin-contain
40	23	15.9	8	2 B45800	serum albumin - do
41	23	15.9	10	2 H37196	bradykinin-potent
42	23	15.9	14	1 NTKN1M	alpha-conotoxin MI
43	23	15.9	14	2 C35141	T-cell receptor de
44	23	15.9	15	2 PHI788	T-cell receptor al
45	23	15.9	15	2 PX0031	mixed lymphocyte r

ALIGNMENTS

RESULT 1

S48156
alpha-amylase inhibitor - rye
C:Species: Secale cereale (rye)
C:Date: 07-May-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S48156
Eur. J. Biochem. 224, 525-531, 1994
R:Garcia-Casado, G.; Sanchez-Monge, R.; Lopez-Otin, C.; Salcedo, G.
A:Title: Rye inhibitors of animal alpha-amylases show different specificities, aggregat
A:Reference number: S48156; MUID:95010030
A:Accession: S48156
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <GAR>
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: alpha-amylase inhibitor

Query Match 22.8%; Score 33; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPFPKPIVRSI 22

DB 3 GEWCWPGMGHPMPFPRCRAL 23

RESULT 2

S26664
microtubule-associated protein tau - human

C:Species: Homo sapiens (man)

C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 26-May-1995

C:Accession: S26664

R:Andreadis, A.; Brown, W.M.; Kosik, K.S.

Biochemistry 31, 10626-10633, 1992

A:Title: Structure and novel exons of the human tau gene.

A:Reference number: S26662; MUID:93041757

A:Accession: S26664

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-18 <AND>

A:Cross-references: EMBL:X61375

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991

Query Match 20.0%; Score 29; DB 2; Length 18;

Best Local Similarity 71.4%; Pred. No. 6.3e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSRVDHP 12

DB 4 KSRDHP 10

A: Experimental source: skin
C: Superfamily: unassigned animal peptides

Query Match 18.6%; Score 27; DB 2; Length 13;
Best Local Similarity 44.4%; Pred. No. 8.8e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 DHPFPKPPI 18
: | : | |
Db 2 EKPFPYPPPI 10

RESULT 9

S09363
hypothetical MTCOL/MTCYB mutant fusion protein .. human mitochondrion (fragment)
C: Species: mitochondrion Homo sapiens (man)
C: Date: 21-Nov-1993 #sequence_revision 14-Aug-1997 #text_change 20-Apr-2000
C: Accession: S09363
R: Poulton, J.; Deadman, M.E.; Gardiner, R.M.
Nucleic Acids Res. 17, 10223-10229, 1989
A: Title: Tandem direct duplications of mitochondrial DNA in mitochondrial myopathy: analysis of a family with a 1.7 kb deletion
A: Reference number: S09363; MUID: 90098864
A: Molecule type: DNA
A: Residues: 1-24 <POU>
C: Comment: This is the hypothetical translation of a sequence believed to result from a
C: Genetics:
A: Gene: MTCOL/MTCYB
A: Genome: mitochondrion
A: Genetic code: SGC1
C: Keywords: fusion protein; mitochondrion
F: 1-4/Region: cytochrome-c oxidase chain I
F: 5-24/Region: cytochrome b (+2 frame shifted)

Query Match 18.6%; Score 27; DB 4; Length 24;
Best Local Similarity 31.2%; Pred. No. 1.7e+03;
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 10 DHPFPKPPIVRSITKC 25
: | : | : | :
Db 6 DHPSTQKPETSALSSC 21

RESULT 10

S09025
carboxylesterase (EC 3.1.1.1), microsomal - rabbit (fragment)
C: Species: Oryctolagus cuniculus (domestic rabbit)
C: Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998
C: Accession: S09025
R: Hosokawa, M.; Maki, T.; Satoh, T.
Arch. Biochem. Biophys. 277, 219-227, 1990
A: Title: Characterization of molecular species of liver microsomal carboxylesterases of
A: Reference number: S09021; MUID: 90179180
A: Accession: S09025
A: Molecule type: protein
A: Residues: 1-20 <HOS>
C: Superfamily: cholinesterase; cholinesterase homology
C: Keywords: carboxylic ester hydrolase

Query Match 18.3%; Score 26.5; DB 2; Length 20;
Best Local Similarity 41.7%; Pred. No. 1.6e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 11 HPFPKPPIVRSI 22
: | : | : | :
Db 1 HPSP-PVNVXV 11

RESULT 11

I78841

thrombopoietin receptor - mouse (fragment)
C: Species: Mus sp. (mouse)
C: Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C: Accession: I78841
R: Alexander, W.S.; Dunn, A.R.
Oncogene 10, 795-803, 1995

A: Title: Structure and transcription of the genomic locus encoding murine c-mpl, a re
A: Reference number: I58350; MUID: 95166571
A: Accession: I78841
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: mRNA
A: Residues: 1-18 <RES>
A: Cross-references: GB:S76842; NID:g912990; PIDN:AAB33462.1; PID:g912991
C: Genetics:
A: Gene: c-mpl1

Query Match 17.9%; Score 26; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 PHFPKPPIV 19
: | : | : | :
Db 2 PHGPAPFL 9

RESULT 12

A41439
acid ribonuclease (EC 3.1.1.-) - bovine (fragment)
C: Species: Bos primigenius taurus (cattle)
C: Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 21-Mar-1996
C: Accession: A41439
R: Ohgi, K.; Senda, A.; Takizawa, Y.; Irie, M.
J. Biochem. 103, 267-273, 1988
A: Title: Purification of acid ribonucleases from bovine spleen.
A: Reference number: A41439; MUID: 88227899
A: Accession: A41439
A: Status: preliminary
A: Molecule type: protein
A: Residues: 1-20 <OHG>
C: Keywords: hydrolase

Query Match 17.9%; Score 26; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPFPKPPI 18
: | : | : | :
Db 9 HLPFPKDL 16

RESULT 13

JN0910
N4-(beta-N-acetylglucosaminyl)-L-asparaginase (EC 3.5.1.26) alpha chain - Flavobacter
N: Alternate names: glycosylasparaginase; N4-(N-acetyl-beta-glucosaminyl)-L-asparagin
C: Species: Flavobacterium meningosepticum
C: Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
C: Accession: JN0910
R: Tarentino, A.L.; Plummer Jr., T.H.
Biochem. Biophys. Res. Commun. 197, 179-186, 1993
A: Title: The first demonstration of a procaryotic glycosylasparaginase.
A: Reference number: JN0910; MUID: 94071939
A: Accession: JN0910
A: Molecule type: protein
A: Residues: 1-22 <TAR>
C: Comment: This heterodimeric enzyme is the counterpart to a lysosomal amidase/amidoh
gine-linked glycans.
C: Keywords: heterodimer; hydrolase

Query Match 17.9%; Score 26; DB 2; Length 22;
Best Local Similarity 75.0%; Pred. No. 2.2e+03;

Matches	6;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
QY	16	KPIVRSIT	23						
Db	3	KPIVLSTT	10						
RESULT 14									
A47415									
mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) 37K beta chain - pig (fragment)									
N:Alternate names: GDP-mannose pyrophosphorylase 37K beta chain									
C:Species: Sus scrofa domestica (domestic pig)									
C:Date: 23-Feb-1994 #sequence_revision 12-Aug-1996 #text_change 13-Sep-1998									
C:Accession: A47415									
R:Szumilo, T.; Drake, R.R.; York, J.L.; Elbein, A.D.									
J. Biol. Chem. 268, 17943-17950, 1993									
A:Title: GDP-mannose pyrophosphorylase. Purification to homogeneity, properties, and utilization									
A:Reference number: A47415; MUID:93352609									
A:Contents: liver									
A:Accession: A47415									
A:Molecule type: protein									
A:Residues: 1-23 <SZU>									
A:Note: sequence extracted from NCBI backbone (NCBIP:136438)									
C:Complex: The enzyme appears to be a heterodimer of alpha and beta chains.									
C:Function:									
A:Description: generates GDP-mannose and pyrophosphate from mannose-1-phosphate and GTP									
A:Note: also catalyzes synthesis of GDP-glucose from glucose-1-phosphate (EC 2.7.7.34 and EC 2.7.7.35)									
C:Superfamily: mannose-1-phosphate guanylyltransferase									
C:Keywords: nucleotidyltransferase									
Query Match 17.9%; Score 26; DB 2; Length 23;									
Best Local Similarity 33.3%; Pred. No. 2.3e+03;									
Matches	7;	Conservative	4;	Mismatches	4;	Indels	6;	Gaps	1;
QY	4	GYSRVDPHPKPPIVRSITK	24						
Db	9	GYGTRL-----RPLTSLIPK	23						
RESULT 15									
S29749									
serum albumin - dog									
C:Species: Canis lupus familiaris (dog)									
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999									
C:Accession: S29749									
R:Dixon, J.W.; Sarkar, B.									
J. Biol. Chem. 249, 5872-5877, 1974									
A:Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide (1-24) from									
A:Reference number: S29749; MUID:75011422									
A:Accession: S29749									
A>Status: preliminary									
A:Molecule type: protein									
A:Residues: 1-24 <DIX>									
C:Superfamily: serum albumin; serum albumin repeat homology									
Query Match 17.9%; Score 26; DB 2; Length 24;									
Best Local Similarity 44.4%; Pred. No. 2.4e+03;									
Matches	4;	Conservative	2;	Mismatches	3;	Indels	0;	Gaps	0;
QY	3	EGYQSRVDH	11						
Db	1	EAYKSEIAH	9						

Search completed: August 26, 2002, 10:20:42

Job time: 348 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:30:17 ; Search time 22.95 Seconds
(without alignments)
42.178 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 1503

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	19.3	13	1 TP13_PHYRO	P04096 phyllomedusa
2	28	19.3	23	1 CP23_SPOER	P56683 spodoptera
3	28	19.3	23	1 PAPI_SPOEX	P30255 spodoptera
4	28	19.3	23	1 PAP2_SPOEX	P30256 spodoptera
5	28	19.3	23	1 PAP3_SPOEX	P30257 spodoptera
6	25.5	17.6	16	1 MK2B_PALPR	P80410 palomena pr
7	24	16.6	10	1 FAR6_PANRE	P82660 panagrellus
8	24	16.6	10	1 MCA2_RHOOP	P56870 rhodococcus
9	24	16.6	24	1 IMPI_GALME	P82176 gallieria me
10	23.5	16.2	15	1 MK1_PALPR	P80408 palomena pr
11	23.5	16.2	15	1 MK2A_PALPR	P80409 palomena pr
12	23.5	16.2	16	1 MK3_PALPR	P80411 palomena pr
13	23	15.9	9	1 NEUX_HUMAN	P04277 homo sapien
14	23	15.9	10	1 BPP8_BOTIN	P30426 bothrops in
15	23	15.9	14	1 CXA1_CONMA	P01521 conus magus
16	23	15.9	19	1 HI70_RAT	P21794 rattus norv
17	23	15.9	20	1 PYRR_PPRAP	P37362 pyrrhocoris
18	23	15.9	23	1 PAPI_MANSE	P30253 manduca sex
19	22.5	15.5	21	1 THAN_PODMA	P55788 podimus mac
20	22.5	15.5	22	1 TX12_TRIWA	P24335 trimeresur
21	22.5	15.5	25	1 CXA4_CONPU	P55963 conus purpu
22	22	15.2	10	1 BPP2_BOTIN	P30422 bothrops in
23	22	15.2	19	1 AMY_DERPT	P49274 dermatophag
24	22	15.2	21	1 CFPA_TREPH	P56738 treponema p
25	22	15.2	21	1 LPRM_CORDI	P21232 corynebacte
26	22	15.2	24	1 BRIA_RANES	P40835 rana escul
27	22	15.2	25	1 H2B1_ECHES	P13281 echinus esc
28	21.5	14.8	15	1 MAOX_CHICK	Q92060 gallus gall
29	21	14.5	13	1 MLA_ANOCA	P41589 anolis caro
30	21	14.5	13	1 MLA_CAMDR	P01198 camelus dro
31	21	14.5	14	1 MY14_EISFO	P46979 eisenia foe
32	21	14.5	14	1 MY14_PHEVI	P46980 pheretima v
33	21	14.5	15	1 ACEA_ACICA	P28467 acinetobact

ALIGNMENTS

```

RESULT 1
TP13_PHYRO
ID TP13_PHYRO STANDARD; PRT; 13 AA.
AC P04096;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 01-NOV-2002 (Rel. 41, Last annotation update)
DE Tryptophyllin-13.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
RP SEQUENCE.
RA Montecucchi P.C., Gozzini L., Erspamer V.;
RT "Primary structure determination of a tryptophan-containing
RT tridecapeptide from Phyllomedusa rohdei."
RL Int. J. Pept. Protein Res. 27:175-182(1986).
CC -!- TISSUE SPECIFICITY: SECRETED.
DR PIR; A05174; A05174.
KW Amphibian skin.
FT MOD_RES 1
SQ SEQUENCE 13 AA; 1646 MW; 33BF33A212227773 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 13;
Best Local Similarity 44.4%; Pred. No. 2.5e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 DHPHFKPI 18
   :|:|:|
Db 2 EKPWPPPI 10

RESULT 2
CP23_SPOER
ID CP23_SPOER STANDARD; PRT; 23 AA.
AC P56683;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Cardioactive peptide CAP23.
OS Spodoptera eridania (Southern armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=37547;
RN [1]
RP SEQUENCE.
RX MEDLINE=99196260; PubMed=10098624;
RA Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H.,
RA Shabanowitz J., Hunt D.F., Schooley D.A.;
RT "A cardioactive peptide from the southern armyworm, Spodoptera
RT eridania."
RL Peptides 20:53-61(1999).

```

Q02210 pseudomonas
P28879 conus stria
P35451 bos taurus
P30251 heliothis v
P30252 heliothis v
P37154 felis silve
P12358 triticum ae
P81050 mesocricetu
P02807 neosporosa
P56973 conus conso
P01519 conus geogr
P19920 pseudomonas

```

CC -!- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM
CC LARVAL MANDUCA SEXTA, CAUSING AN INOTROPIC EFFECT AT LOW
CC CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS
CC AT HIGH DOSES.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR HSSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2519 MW; 0A96D72A70855AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVD 10
DB 7 CTFGYQRTAD 16

RESULT 3
PAP1_SPOEX
ID PAP1_SPOEX STANDARD; PRT; 23 AA.
AC P30255;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Paralytic peptide I (pp I).
OS Spodoptera exigua (Beet armyworm).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
CC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens.";
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; C39855; C39855.
DR HSSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2451 MW; 0A96D1F600855AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVD 10
DB 7 CTFGYQRTAD 16

RESULT 4
PAP2_SPOEX
ID PAP2_SPOEX STANDARD; PRT; 23 AA.
AC P30256;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

```

DE Paralytic peptide II (pp II).
OS Spodoptera exigua (Beet armyworm).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
CC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens.";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; C39855; D39855.
DR HSSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2477 MW; 0A96CB4600855AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVD 10
DB 7 CTFGYQRTAD 16

RESULT 5
PAP3_SPOEX
ID PAP3_SPOEX STANDARD; PRT; 23 AA.
AC P30257;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Paralytic peptide III (pp III).
OS Spodoptera exigua (Beet armyworm).
CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
CC Noctuoidea; Noctuidae; Amphipyridae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens.";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -!- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -!- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; C39855; E39855.
DR HSSP; 061704; 1BSN.
DR InterPro; IPR003463; GBP_PSP.
DR Pfam; PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2505 MW; 0A96CB5EB7D55AE0 CRC64;

```

```
Query Match      19.3%; Score 28; DB 1; Length 23;
Best Local Similarity 50.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

Qy 1 CGEGYQSRVD 10
Db 7 CTPGYORTAD 16

RESULT	6
MK2B_PALPR	
ID	MK2B_PALPR STANDARD; PRG; 16 AA.
AC	P80410;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update);
DT	15-JUL-1998 (Rel. 36, Last annotation update)
DE	Metainkikowin IIB.
OS	Palomena prasina.
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC	Pterygota; Neoptera; Paransoptera; Hemiptera; Euhemiptera;
OC	Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC	Pentatomidae; Palomena.
OX	NCBI_Taxid=55431;
RN	[1]
RP	SEQUENCE.
RC	TISSUE=Hemolymph;
RA	Chernysh S., Cocciandich S., Briand J.-P., Hetru C., Bulet P.;
RT	"The inducible antibacterial peptides of the hemipteran insect
RT	Palomena prasina. Identification of a unique family of proline-rich
RT	peptides and of a novel insect defensin.";
RL	J. Insect Physiol. 42:81-89(1996).
CC	-!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC	BACTERIA.
CC	-!- INDUCTION: BY BACTERIAL INFECTION.
KW	Antibiotic; Insect immunity.
SQ	SEQUENCE 16 AA; 2040 MW; AF21407D063B9462 CRC64;

Query Match 17.6%; Score 25.5; DB 1; Length 16;
Best Local Similarity 40.0%; Pred. No. 7.21+02;
Matches 6; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 9 V D H P H F - P K P I V R S I 22
||| : | : | : :
db 1 V D K P D Y R P R P W P R N M 15

RESULT	7
FAR6_PANRE	
ID	FAR6_PANRE
AC	PRT;
AC	STANDARD; PRT; 10 AA.
AC	P82660;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	FMRFamide-like neuropeptide PF6 (NGAPQPFVR?-amide).
OS	Panagrellus redivivus.
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OX	Panagrolaimoidea; Panagrolaimidae; Panagrellus.
NCBI_TaxID=6233;	
[1]	
RN	SEQUENCE, FUNCTION, AND AMIDATION.
RP	Moffett C.L., Marks N.J., Halton D.W., Thomson D.P., Geary T.G.,
RA	Maule A.G.;
RT	"Isolation, characterization and pharmacology of FMRFamide-related peptides (fARPs) from free-living nematode, Panagrellus redivivus."
RL	Submitted (JUL-2000) to the SWISS-PROT data bank.
CC	-!- FUNCTION: MYOACTIVE.
CC	-!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE) FAMILY.
KW	Neuropeptide; Amidation.
MOD_RES	10
FT	SEQUENCE 10 AA; 1132 MW; CB13E4C9D776C76D CRC64;
SQ	

Query Match 16.6%; Score 24; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 7.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels

Qy 15 PKPIVR 20
|:| |
Db 4 POPFVR 9

RESULT	8			
MC_A2	RHOOP			
ID	MC_A2_RHOOP	STANDARD;	PRT;	15 AA.
AC	P56870;			
DT	30-MAY-2000	(Rel. 39, Created)		
DT	30-MAY-2000	(Rel. 39, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Putative maleylacetate reductase II	(EC 1.3.1.32) (Fragment).		
OS	Rhodococcus opacus	(Nocardia opaca).		
OC	Bacteria; Firmicutes; Actinobacteria;	Actinobacteridae;		
OC	Actinomycetales; Corynebacterineae;	Nocardiaceae; Rhodococcus.		
OX	NCBI_TaxID=37919;			
RN	{1			
RC	SEQUENCE.			
RC	STRAIN=LCP;			
RX	MEDLINE=98324954; PubMed=9657989;			
RA	Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;			
RT	"Characterization of a maleylacetate reductase encoding region from			
RT	Rhodococcus opacus LCP.";			
RL	J. Bacteriol. 180:3503-3508(1998).			
CC	-1- CATALYTIC ACTIVITY: 3-oxoadipate + NAD(P)(+) -> 2-maleylacetate +			
CC	NAD(P)H.			
CC	-1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY)			
CC	THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC			
CC	AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL			
CC	PRODUCTS AND AS INDUSTRIAL EFFLUENT.			
CC	-1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE			
CC	FAMILY.			

```
Query Match          16.6%; Score 24; DB 1; Length 15;
Best Local Similarity 36.4%; Pred. No. 1.1e+03;
Matches 4: Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

Qy 8 RVDHPHPKPI 18
| : | : | : |
Db 2 RFEHENLPORI 12

RESULT	9	IMPI_GALME	IMPI_GALME	STANDARD;	PRT;	24 AA.
AC		P82176;				
DT		16-OCT-2001	(Rel. 40, Created)			
DT		16-OCT-2001	(Rel. 40, Last sequence update)			
DT		16-OCT-2001	(Rel. 40, Last annotation update)			
DE		Metalloproteinase inhibitor (IMPI) (fragment).				
OS		Galleria mellonella (Wax moth).				
OS		Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC		Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;				
OC		Pyraloidea; Pyralidae; Galleriinae; Galleria.				
OX		NCBI_TaxID=7137;				
RN		[1]				
RP		SEQUENCE.				
RT		TISSUE=Larval hemolymph;				
RX		MEDLINE=98409271; PubMed=9738891;				

RA Wedde M., Weise C., Kopacek P., Franke P., Vilcinskis A.;
 RT "Purification and characterization of an inducible metalloprotease
 RT inhibitor from the hemolymph of greater wax moth larvae, *Galleria*
 RT *mellonella*.";
 RL Eur. J. Biochem. 255:535-543(1998).
 CC -|- FUNCTION: IT INHIBITS THERMOLYSIN AND THE METALLOPROTEASE OF
 CC B. POLYMYXA. NO ACTIVITY ON TRYPSIN OR CYSTEINPROTEASE PAPAN.
 CC -|- INDUCTION: BY LIPOPOLYSACCHARIDE (LPS).
 CC -|- PTM: FIVE DISULFIDE BONDS ARE PRESENT.
 CC -|- PTM: THE HEAVY CHAIN MAY BE O-GLYCOSYLATED.
 CC -|- MASS SPECTROMETRY: MW=8360; METHOD=MALDI.
 KW Metalloprotease inhibitor; Glycoprotein.
 FT NON_TER 1 1
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2783 MW; 490E278E449F0C25 CRC64;

Query Match 16.6%; Score 24; DB 1; Length 24;
 Best Local Similarity 33.3%; Pred. No. 1.8e+03;
 Matches 5; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGEGYOSRVDHPHF 15
 Db 9 CEDGYARDVNGKCIP 23

RESULT 10
 MK1_PALPR STANDARD; PRT; 15 AA.
 AC P80408;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Metalnikowin I.
 OS Palomena prasina.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomioidea;
 OC Pentatomidae; Palomena.
 OX NCBI_TaxID=55431;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
 RT "The inducible antibacterial peptides of the hemipteran insect
 RT *Palomena prasina*. Identification of a unique family of proline-rich
 RT peptides and of a novel insect defensin.";
 RL J. Insect Physiol. 42:81-89(1996).
 CC -|- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -|- INDUCTION: BY BACTERIAL INFECTION.
 KW Antibiotic; Insect immunity.
 SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

Query Match 16.2%; Score 23.5; DB 1; Length 15;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 VDPHPF-PKP 17
 Db 1 VDKPDYRPRP 10

RESULT 11
 MK2A_PALPR STANDARD; PRT; 15 AA.
 AC P80409;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Metalnikowin IIA.
 OS Palomena prasina.

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomioidea;
 OC Pentatomidae; Palomena.
 OX NCBI_TaxID=55431;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
 RT "The inducible antibacterial peptides of the hemipteran insect
 RT *Palomena prasina*. Identification of a unique family of proline-rich
 RT peptides and of a novel insect defensin.";
 RL J. Insect Physiol. 42:81-89(1996).
 CC -|- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -|- INDUCTION: BY BACTERIAL INFECTION.
 KW Antibiotic; Insect immunity.
 SQ SEQUENCE 15 AA; 1893 MW; 23835D063B946299 CRC64;

Query Match 16.2%; Score 23.5; DB 1; Length 15;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 VDPHPF-PKP 17
 Db 1 VDKPDYRPRP 10

RESULT 12
 MK3_PALPR STANDARD; PRT; 16 AA.
 AC P80411;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Metalnikowin III.
 OS Palomena prasina.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomioidea;
 OC Pentatomidae; Palomena.
 OX NCBI_TaxID=55431;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
 RT "The inducible antibacterial peptides of the hemipteran insect
 RT *Palomena prasina*. Identification of a unique family of proline-rich
 RT peptides and of a novel insect defensin.";
 RL J. Insect Physiol. 42:81-89(1996).
 CC -|- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -|- INDUCTION: BY BACTERIAL INFECTION.
 KW Antibiotic; Insect immunity.
 SQ SEQUENCE 16 AA; 2024 MW; A9E3835D063B9462 CRC64;

Query Match 16.2%; Score 23.5; DB 1; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.4e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 VDPHPF-PKP 17
 Db 1 VDKPDYRPRP 10

RESULT 13
 NEUX_HUMAN STANDARD; PRT; 9 AA.
 AC P04277;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)

01-MAR-2002 (Rel. 41, Last annotation update)
 DE Neurotensin-related peptide (NRP) (Kinetensin).
 OS Homo sapiens (Human), and
 OS Bos taurus (Bovine), and
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606, 9913, 9986;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Human;
 RX MEDLINE=86242180; PubMed=3087352;
 RA Mogard M.H., Kobayashi R., Chen C.F., Lee I.D., Reeve J.R. Jr.,
 RA Shively J.E., Walsh J.H.;
 RT "The amino acid sequence of kinetensin, a novel peptide isolated from
 RT pepsin-treated human plasma: homology with human serum albumin,
 RT neurotensin and angiotensin.";
 RL Biochem. Biophys. Res. Commun. 136:983-988(1986).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=Human, Bovine, and Rabbit;
 RX MEDLINE=87194805; PubMed=2437111;
 RA Carraway R.E., Mitra S.P., Cochrane D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 RT obtained from pepsin-treated albumin(s).";
 RL J. Biol. Chem. 262:5968-5973(1987).
 CC -|- FUNCTION: REGULATION OF FAT DIGESTION, LIPID ABSORPTION, AND
 CC BLOOD FLOW (POTENTIAL).
 DR PIR; A03239; ABHUSK.
 KW Hormone.
 SQ SEQUENCE 9 AA; 1172 MW; C804DB4761F4140D CRC64;

Query Match 15.9%; Score 23; DB 1; Length 9;
 Best Local Similarity 75.0%; Pred. No. 1e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPHF 14
 DB 5 HPYF 8

RESULT 14
 BPP8_BOTIN STANDARD; PRT; 10 AA.
 AC P30426;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Bradykinin-potentiating peptide S5,1 (Angiotensin-converting
 DE enzyme inhibitor).
 OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8723;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=90351557; PubMed=2386615;
 RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
 RT "Primary structure and biological activity of bradykinin potentiating
 RT peptides from Bothrops insularis snake venom.";
 RL J. Protein Chem. 9:221-227(1990).
 CC -|- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
 DR PIR; H37196; H37196.
 KW Hypotensive agent; Venom.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 SQ SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 1e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHP 15
 DB 5 HPNP 9

RESULT 15
 CXAL_CONMA STANDARD; PRT; 14 AA.
 AC P01521;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alpha-conotoxin MI (M1).
 OS Conus magus (Magus cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6492;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=83073458; PubMed=7149738;
 RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
 RT "Isolation and structure of a peptide toxin from the marine snail
 RT Conus magus.";
 RL Arch. Biochem. Biophys. 218:329-334(1982).
 RN [2]
 RP DISULFIDE BONDS.
 RX MEDLINE=84032400; PubMed=6630187;
 RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
 RT "Conotoxin MI. Disulfide bonding and conformational states.";
 RL J. Biol. Chem. 258:12247-12251(1983).
 RN [3]
 RP REVIEW.
 RX MEDLINE=89024586; PubMed=3052286;
 RA Gray W.R., Olivera B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Annu. Rev. Biochem. 57:665-700(1988).
 CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 CC INHIBIT THEM.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
 DR PIR; A01784; NTKNIM.
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
 KW Venom.
 FT DISULFID 3 8
 FT DISULFID 4 14
 FT MOD_RES 14 14 AMIDATION.
 SQ SEQUENCE 14 AA; 1499 MW; DEEE91898BF5E5BD CRC64;

Query Match 15.9%; Score 23; DB 1; Length 14;
 Best Local Similarity 60.0%; Pred. No. 1.4e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
 DB 8 CGKNY 12

Search completed: August 26, 2002, 10:30:18
 Job time: 638 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:29:34 ; Search time 62.22 Seconds
(without alignments)
69.509 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CEEGYQSRVDHPFKPIVRSITKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 8561

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	32	22.1	22	4 Q13659	Q13659 homo sapien
2	30	20.7	17	4 Q16310	Q16310 homo sapien
3	30	20.7	18	2 Q56610	Q56610 vibrio chol
4	30	20.7	19	13 P87468	P87468 xenopus lae
5	30	20.7	22	11 Q9QV92	Q9QV92 rattus sp.
6	29	20.0	20	2 Q52318	Q52318 rhizobium m
7	28.5	19.7	23	12 Q9QSP0	Q9QSP0 australian
8	28	19.3	14	4 Q9P2A2	Q9P2A2 homo sapien
9	28	19.3	21	2 Q44041	Q44041 anabaena fl
10	28	19.3	23	15 Q85496	Q85496 bovine leuk
11	28	19.3	25	4 Q9UKK1	Q9UKK1 homo sapien
12	28	19.3	25	5 Q9BK01	Q9BK01 hydra litto
13	28	19.3	25	6 Q9TRY2	Q9TRY2 sus sp. ins
14	27.5	19.0	15	15 Q905H2	Q905H2 human immun
15	27	18.6	11	5 Q9NL65	Q9NL65 ascaris suu
16	27	18.6	15	4 Q9UCC0	Q9UCC0 homo sapien

17	27	18.6	17	2	Q34216	Q34216 sphingomona
18	26	17.9	17	6	Q9TR98	Q9TR98 canis famil
19	26	17.9	22	7	Q9MX47	Q9MX47 oryzias lat
20	26	17.9	22	11	Q35905	Q35905 mus musculu
21	26	17.9	23	6	Q9TRF4	Q9TRF4 sus scrofa
22	26	17.9	24	3	Q9UR88	Q9UR88 aspergillus
23	26	17.9	24	15	Q9TNP5	Q9TNP5 human immun
24	26	17.9	25	4	Q16092	Q16092 homo sapien
25	25	17.2	15	6	Q9TQ09	Q9TQ09 bos taurus
26	25	17.2	15	6	Q9TR40	Q9TR40 bos taurus
27	25	17.2	16	2	Q9LAP2	Q9LAP2 enterococcu
28	25	17.2	16	4	Q16350	Q16350 homo sapien
29	25	17.2	17	2	Q9WW21	Q9WW21 enterococcu
30	25	17.2	17	2	Q9WW20	Q9WW20 enterococcu
31	25	17.2	17	2	Q9X5I5	Q9X5I5 enterococcu
32	25	17.2	18	6	Q95JA2	Q95JA2 sus scrofa
33	25	17.2	19	2	Q33755	Q33755 streptococc
34	25	17.2	19	10	O80997	O80997 arabidopsis
35	25	17.2	20	11	Q61871	Q61871 mus musculu
36	25	17.2	21	4	Q9UCC5	Q9UCC5 homo sapien
37	25	17.2	21	4	Q16017	Q16017 homo sapien
38	25	17.2	21	12	Q93191	Q93191 porcine cir
39	25	17.2	21	12	Q9YQS8	Q9YQS8 porcine cir
40	25	17.2	21	12	O56126	O56126 porcine cir
41	25	17.2	22	4	Q96Q47	Q96Q47 homo sapien
42	25	17.2	23	12	Q65291	Q65291 human adeno
43	25	17.2	24	2	Q9R5H2	Q9R5H2 lactobacill
44	25	17.2	25	4	Q16498	Q16498 homo sapien
45	24	16.6	16	8	Q34374	Q34374 daphnia pul

ALIGNMENTS

RESULT 1

Q13659	ID	Q13659	PRELIMINARY;	PRT;	22 AA.
AC	Q13659;				
DT	01-NOV-1996	(TrEMBLrel. 01, Created)			
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)			
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)			
DE	CARDIAC MYOSIN BINDING PROTEIN-C (FRAGMENT).				
GN	MYBP-C.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96083593; PubMed=7493026;				
RA	Bonne G., Carrier L., Bercovici J., Cruaud C., Richard P., Hainque B.,				
RA	Gautel M., Labelit S., James M., Beckmann J., et al;				
RT	"Cardiac myosin binding protein-C gene splice acceptor site mutation				
RL	is associated with familial hypertrophic cardiomyopathy.";				
RL	Nat. Genet. 11:438-440(1995).				
DR	EMBL; S80782; AAB35663.1; -;				
FT	NON_TER				
SQ	SEQUENCE	22 AA;	2563 MW;	5BA92048F76774CA	CRC64;

Query Match 22.1% Score 32; DB 4; Length 22;
Best Local Similarity 38.9%; Pred. No. 3e+02;
Matches 7; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 8 RVDHPFPKPIVRSITKC 25
| | | | | | | | | |
Db 2 RQDPGLPRPHRHGCG 19

RESULT 2
Q16310
ID Q16310
AC Q16310; PRELIMINARY; PRT; 17 AA.

```
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE GC*2 PROTEIN (FRAGMENT).
GN GC*2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95242701; PubMed=7725672;
RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;
RT "Characterization of mutants of the vitamin-D-binding protein/group
RT specific component: GC aborigine (Ia1) from Australian aborigines and
RT South African blacks, and 2A9 from south Germany.";
RL Vox Sang. 68:50-54(1995).
DR EMBL; S77130; RAD14250.1; -.
DR NON_TER 17
FT SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;
SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 20.7%; Score 30; DB 4; Length 17;
Best Local Similarity 36.4%; Pred. No. 4.8e+02;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 15 PKPIVRSITKC 25
   ||| : : ||
Db 6 PKELAKLVNKC 16

RESULT 3
Q56610 Q56610 PRELIMINARY; PRT; 18 AA.
AC Q56610;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ACCA (FRAGMENT).
GN ACCA.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C6706;
RX MEDLINE=97074686; PubMed=8917113;
RA Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,
RA Morris J.G.;
RT "Cloning and characterization of dnaE, encoding the catalytic subunit
RT of replicative DNA polymerase III, from Vibrio cholerae strain
RT C6706.";
RL Gene 175:281-283(1996).
DR EMBL; U30472; AAC44579.1; -.
DR NON_TER 18
FT SEQUENCE 18 AA; 2153 MW; 18EBBCDAD212842EF CRC64;
SQ SEQUENCE 18 AA; 2153 MW; 18EBBCDAD212842EF CRC64;

Query Match 20.7%; Score 30; DB 2; Length 18;
Best Local Similarity 53.6%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 14 FPKPIVRSITK 24
   ||||| ||
Db 8 FEPKPIVELETK 18

RESULT 4
P87468 P87468 PRELIMINARY; PRT; 19 AA.
ID P87468
AC P87468;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALPHA-GLOBIN (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81001900; PubMed=6250724;
RA Patient R.K., Elkington J.A., Kay R.M., Williams J.G.;
RT "Internal organization of the major adult alpha- and beta-globin genes
RT of X.laevis.";
RL Cell 21:565-573(1980).
CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
DR EMBL; V01431; CAA24695.1; -.
DR InterPro; IPR000971; Globin.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Respiratory protein.
FT NON_TER 1
FT SEQUENCE 19 AA; 2144 MW; AAC19D73A95C9AC6 CRC64;
SQ SEQUENCE 19 AA; 2144 MW; AAC19D73A95C9AC6 CRC64;

Query Match 20.7%; Score 30; DB 13; Length 19;
Best Local Similarity 54.5%; Pred. No. 5.4e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 YQSRVDHPHP 15
   | ||| : ||
Db 9 YLVRVDPGNFP 19

RESULT 5
Q9QV92 Q9QV92 PRELIMINARY; PRT; 22 AA.
ID Q9QV92
AC Q9QV92;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE P93 CALCIUM-BINDING PROTEIN (FRAGMENT).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP SEQUENCE.
RX MEDLINE=93179436; PubMed=8440713;
RA Glchrist J.S., Pierce G.N.;
RT "Identification and purification of a calcium-binding protein in
RT hepatic nuclear membranes.";
RL J. Biol. Chem. 268:4291-4299(1993).
DR InterPro; IPR001580; Calreticulin.
DR Pfam; PF00262; Calreticulin; 1.
DR ProDom; PD001866; Calreticulin; 1.
DR NON_TER 22
FT SEQUENCE 22 AA; 2407 MW; B82830E4477E25F9 CRC64;
SQ SEQUENCE 22 AA; 2407 MW; B82830E4477E25F9 CRC64;

Query Match 20.7%; Score 30; DB 11; Length 22;
Best Local Similarity 28.6%; Pred. No. 6.4e+02;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHF 14
   || : : || : :
Db 7 CGVQRPMIDPNV 20

RESULT 6
Q52918 Q52918 PRELIMINARY; PRT; 20 AA.
ID Q52918
AC Q52918;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

DE HYPOHETICAL 2.4 KDA PROTEIN.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=65297790; PubMed=2994020;
RA Leong S.A., Williams P.H., Ditta G.S.;
RT "Analysis of the 5' regulatory region of the gene for delta-
aminolevulinic acid synthetase of Rhizobium meliloti.";
RL Nucleic Acids Res. 13:5965-5976(1985).
DR EMBL: X02853; CAA26608.1; -.
KW Hypothetical protein.
SQ SEQUENCE 20 AA; 2441 MW; BE09CF11ABCC3BB2 CRC64;

Query Match 20.0%; Score 29; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 8.3e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 HFPKPIV 19
 | | | | :
DB 4 HFPSPFL 10

RESULT 7
Q9QSP0 PRELIMINARY; PRT; 23 AA.
ID Q9QSP0
AC Q9QSP0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE L. PROTEIN (FRAGMENT).
GN L.
OS Australian bat lyssavirus.
OC Viruses; ssRNA negative-strand viruses; Moronegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=90961;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN-INSECTIVOROUS ISOLATE;
RA Gould A.R., Kattenbelt J.A., Hyatt A.D., Gumley S.G., Lunt R.A.;
RT "Characterisation of a variant of Australian Bat Lyssavirus isolated
from an insectivorous bat and comparison to virus isolates from
Pteropid bats.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF081020; AAD47900.1; -.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2599 MW; 7B584A2225E76530 CRC64;

Query Match 19.7%; Score 28.5; DB 12; Length 23;
Best Local Similarity 35.0%; Pred. No. 1.2u+03;
Matches 7; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 2 GEGYQSRVDHPHFKPIVRS 21
 | | | : | | | : :
DB 5 GEVYDDPID-PVEPELKT 23

RESULT 8
Q9P2A2 PRELIMINARY; PRT; 14 AA.
ID Q9P2A2
AC Q9P2A2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TRUNCATED ALDO-KETO REDUCTASE (FRAGMENT).
GN TRUNCATED AKR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20138537; PubMed=10672042;
RA Nishizawa M., Nakajima T., Yasuda K., Kanzaki H., Sasaguri Y.,
RA Watanabe K., Ito S.;
RT "Close kinship of human 20alpha-hydroxysteroid dehydrogenase gene with
three aldo-keto reductase genes.";
RL Genes Cells 5:111-125(2000).
DR EMBL; AB037903; BAA92888.1; -.
FT NON_TER 1
SQ SEQUENCE 14 AA; 1632 MW; 47EB1EE28D59A8D7 CRC64;

Query Match 19.3%; Score 28; DB 4; Length 14;
Best Local Similarity 66.7%; Pred. No. 8.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 HFPKPI 18
 | | | :
DB 2 HFPVPL 7

RESULT 9
Q44041 PRELIMINARY; PRT; 21 AA.
ID Q44041
AC Q44041; Q44040;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUL-1997 (TrEMBLrel. 04, Last annotation update)
DE HYPOHETICAL 2.6 KDA PROTEIN.
OS Anabaena flos-aquae.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1166;
RN [1]
RN SEQUENCE FROM N.A.
RA Hayes P.K.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DDBJ databases.
DR EMBL; M32060; AAA82498.1; -.
DR EMBL; M32060; AAA82496.1; -.
KW Hypothetical protein.
SQ SEQUENCE 21 AA; 2611 MW; 0E4C03BE93E16532 CRC64;

Query Match 19.3%; Score 28; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHFP 15
 | | | |
DB 9 PHFP 12

RESULT 10
Q85496 PRELIMINARY; PRT; 23 AA.
ID Q85496
AC Q85496;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Bovine leukemia virus (BLV).
OC Viruses; Retroviridae; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11901;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=87198886; PubMed=3033284;
RA Rice N.R., Simsek S.L., Dubois G.C., Showalter S.D., Gilden R.V.,
RA Stephens R.M.;
RT "Expression of the bovine leukemia virus X region in virus-infected
cells.";
RL J. Virol. 61:1577-1585(1987).

DR EMBL; M16017; AAA87336.1; -.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2877 MW; F85E7CBED5440B08 CRC64;

Query Match 19.3%; Score 28; DB 15; Length 23;
Best Local Similarity 44.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVRSIT 23
Db 11 POPIIRWVS 19

RESULT 11
Q9UKK1 PRELIMINARY; PRT; 25 AA.
AC Q9UKK1; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE METALLOPROTEINASE-DISINTEGRIN (FRAGMENT).
GN ADAM23
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99453762; PubMed=10524237;
RA Poindexter K., Nelson N., DuBoise R.F., Black R.A., Cerretti D.P.;
RT "The identification of seven metalloproteinase-disintegrin (ADAM)
RT genes from genomic libraries."
RL Gene 237:61-70(1999).
DR EMBL; AF158641; AAD55252.1; -.
KW Integrin.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 2705 MW; 818B03B30EFOA19C CRC64;

Query Match 19.3%; Score 28; DB 4; Length 25;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
Db 7 CGNGY 11

RESULT 12
Q9BM01 PRELIMINARY; PRT; 25 AA.
AC Q9BM01; 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE MARINER-LIKE TRANSPOSASE (FRAGMENT).
OS Hydra littoralis (swiftwater hydra).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=42240;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSOS-MARI;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and asexual taxa."
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY014005; AAG59977.1; -.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 2777 MW; CBF9DD69D7AD074C CRC64;

Query Match 19.3%; Score 28; DB 5; Length 25;
Best Local Similarity 36.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 12 PHFPPIVRSI 22
Db 3 PHVAOPTLRKL 13

RESULT 13
Q9TRY2 PRELIMINARY; PRT; 25 AA.
AC Q9TRY2; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE INSULIN-LIKE GROWTH FACTOR-BINDING PROTEIN-6, IGFBP-6.
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9826;
RN [1]
RP SEQUENCE.
RX MEDLINE=92049376; PubMed=1719383;
RA Shimazaki S., Gao L., Shimonaka M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding
RT protein-6."
RL Mol. Endocrinol. 5:938-948(1991).
DR InterPro; IPR000867; IGFBP.
DR Pfam; PF00219; IGFBP; 1.
SQ SEQUENCE 25 AA; 2310 MW; 4077663151E71212 CRC64;

Query Match 19.3%; Score 28; DB 6; Length 25;
Best Local Similarity 57.1%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGYOS 7
Db 6 CGQGVQT 12

RESULT 14
Q905H2 PRELIMINARY; PRT; 19 AA.
AC Q905H2; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE TAT PROTEIN (FRAGMENT).
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11876;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CG275;
RA Taniguchi Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., M'Pandi M., M'Pele P., Harada Y., Hayami M.,
RA Ichimura H., Farra H.-J.;
RT "Genetic Subtypes of HIV Type 1 Based on the vpu/env Sequences in
RT Republic of Congo-Brazzaville."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410446; AAL10250.1; -.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 2290 MW; 366296E191128236 CRC64;

Query Match 19.0%; Score 27.5; DB 15; Length 19;
Best Local Similarity 58.3%; Pred. No. 1.4e+03;
Matches 7; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 6 QSRVDHPHPKP 17
 ||| || : | |
 Db 8 QSRDHQN-PVP 18

RESULT 15
 Q9NL65
 ID Q9NL65 PRELIMINARY; PRT; 11 AA.
 AC Q9NL65;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE ASABF-DELTA (FRAGMENT).
 GN ASABF-DELTA.
 OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
 OC Ascarididae; Ascaris.
 OX NCBI_TaxID=6253;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kato Y.;
 RT "Ascaris suum asabf-delta gene, exon 2.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AB029815; BAA89496.1; -
 FT NON_TER 1 1
 FT NON_TER 11 11
 SQ SEQUENCE 11 AA; 1187 MW; 8BADDD0CD1EAB5f61 CRC64;

Query Match 18.6%; Score 27; DB 5; Length 11;
 Best Local Similarity 80.0%; Pred. No. 9.2e+02;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
 || ||
 Db 2 CGTGY 6

Search completed: August 26, 2002, 10:29:35
 Job time: 650 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:18:42 ; Search time 79.43 Seconds
(without alignments)
34,960 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameter: 253524

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_032802.*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	100.0	25	21	AA1980000
2	106	73.1	25	21	AA1980001
3	105	72.4	25	21	AA1999999
4	103	71.0	25	21	AA1999998
5	103	71.0	25	21	AA191212
6	103	71.0	25	21	AA198602
7	89	61.4	25	21	AA198077
8	68	46.9	25	18	AAW24102
9	63	43.4	17	21	AA198093
10	63	43.4	17	21	AA198094
11	62	42.8	22	20	AA192585

12	62	42.8	24	20	AA192586	IgE peptide antago
13	58	40.0	17	21	AA198096	Antibody 15A.2 gre
14	55	37.9	17	21	AA198095	Antibody 15A.2 swi
15	54	37.2	17	21	AA198098	IgE derived target
16	51	35.2	25	21	AA198048	Peptide which bind
17	50	34.5	17	11	AA1980179	Antibody 15A.2 mur
18	48	33.1	15	21	AA198099	Antibody 15A.2 fel
19	47	32.4	17	21	AA198097	Human IgE C epsilon
20	44	30.3	19	21	AA1980503	Peptide P7X derive
21	44	30.3	19	22	AA198030	IgE peptide #11.
22	44	30.3	19	22	AA198033	Human IgE C epsilon
23	43	29.7	13	21	AA1980518	Peptide P13b deriv
24	43	29.7	13	22	AA198049	Peptide #3 having
25	41	28.3	22	16	AA198086	Peptide #4 having
26	39	26.9	22	16	AA198087	Human CAMP-specifi
27	37.5	25.9	18	22	AA1980754	Anti-amyloid pepti
28	37	25.5	15	18	AA1980537	Peptide #10204 enc
29	36	24.8	25	22	AA1980589	Human brain expres
30	36	24.8	25	22	AA1980589	Human bone marrow
31	36	24.8	25	22	AA1980404	Peptide #10550 enc
32	36	24.8	25	22	AA1980513	Monoclonal antibod
33	35	24.1	9	21	AA1980599	Monoclonal antibod
34	35	24.1	9	21	AA1980600	Antibody 15A.2 bin
35	35	24.1	9	21	AA1980891	Antibody 15A.2 bin
36	35	24.1	9	21	AA1980892	Antibody 15A.2 bin
37	35	24.1	14	21	AA1980519	Human IgE C epsilon
38	35	24.1	14	21	AA1980579	Immunogenic peptid
39	35	24.1	14	22	AA1980580	Peptide P13c deriv
40	34.5	23.8	12	21	AA1980517	Human IgE C epsilon
41	34.5	23.8	12	22	AA198048	Peptide P13a deriv
42	34.5	23.8	16	21	AA1980495	Human apoptosis re
43	34	23.4	12	21	AA1980516	Human IgE C epsilon
44	34	23.4	12	22	AA198047	Peptide P13 derive
45	34	23.4	12	22	AA1980503	IgE peptide #16.

ALIGNMENTS

RESULT 1

AA1980000

ID AA1980000 standard; Peptide; 25 AA.

AC AA1980000;

XX 15-MAY-2000 (first entry)

DT Optimised IgE-CH3 domain antigen peptide for rat IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Rattus sp.

OS Synthetic.

XX WO9967293-A1.

PN 29-DEC-1999.

PD 21-JUN-1999; 99WO-US13959.

PF 20-JUN-1998; 98US-0100287.

PR (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

PI WPI; 2000-160578/14.

DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX for immunization against allergy

PT

CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 72.4%; Score 105; DB 21; Length 25;

Best Local Similarity 76.0%; Pred. No. 7.3e-09;

Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| | ||| ||| | ||| ||| |||

Db 1 cgetyysrvthphlphkdivrsiack 25

RESULT 4

AAY79998

ID AAY79998 standard; Peptide; 25 AA.

XX AC AAY79998;

XX DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PT for immunization against allergy -

XX PS Claim 1; Page 21; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies
XX CC specific for a target effector site on the epsilon-heavy chain of IgE,
XX CC and so preventing triggering and activation of mast cells and basophils
XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;

Best Local Similarity 68.0%; Pred. No. 1.5e-08;

Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| | ||| ||| | : || |||

Db 1 cgetyysrvthphlpralmrsttkc 25

RESULT 5

AAY91212

ID AAY91212 standard; peptide; 25 AA.

XX AC AAY91212;

XX DT 22-MAY-2000 (first entry)

XX DE Modified human IgE CH3 domain, SEQ ID NO: 92.

XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;

XX KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

XX KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

XX KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; PMDV;

XX KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

XX KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

XX KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9966957-A2.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13975.

XX PR 20-JUN-1998; 98US-0100412.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX WPI; 2000-160564/14.

XX PT New artificial T helper cell epitope and derived immunogens with target
XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX PT or human immune deficiency virus -

XX PS Example 6; Page 40; 129pp; English.

XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
XX CC and immunogenic peptides comprising the Th epitopes of the invention
XX CC along with B cell epitopes. The Th epitopes and peptide immunogens
XX CC containing them, are used to induce a T helper cell response,
XX CC specifically against Plasmodium falciparum, cholesteryl ester transport
XX CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
XX CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX CC peptide immunogens may be used for prevention and/or treatment of
XX CC infections (HIV, foot-and-mouth disease or malaria); for cancer
XX CC immunotherapy; for inhibition of the action of luteinising hormone
XX CC releasing hormone (LHRH) for contraception, treatment of hormone-
XX CC dependent cancer, prevention of boar taint in meat, and
XX CC immunocastration; for promoting the growth of animals; or for
XX CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX CC Th (functional in genetically diverse subjects) into an immunogen
XX CC improves capacity to induce a strong T helper cell-mediated immune
XX CC response, resulting in production of antibodies against a target

antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AAY91121 represents a promiscuous T helper epitope from the measles virus F (MVF) protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AAY90212 is a modified version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a CERP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory invasive protein epitope from *Yersinia* species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

XX Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.5e-08;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPFKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 1 cgetyqsrvtphlpralmrstkc 25

RESULT 6
AAY68602
ID AAY68602 standard; peptide: 25 AA.

AC AAY68602;
DT 05-MAY-2000 (first entry)
DE Peptide sequence of the invention.

XX Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

OS Unidentified.
XX WO9566952-A1.
PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13960.
XX 20-JUN-1998; 98US-0100414.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinising hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer -

PS Disclosure; Page 92; 102pp; English.

XX The specification describes peptide immunogens comprising a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The peptide immunogens cause
CC induction of a specific immune response to LHRH which is involved in
CC regulation of spermatogenesis, ovulation, oestrus, sexual development
CC and secretion of sex hormones. Provision of a promiscuous T helper
CC epitope (which is functional in genetically diverse subjects) provides
CC optimum immunogenicity to the B cell epitopes of the target antigen and
CC thus high antibody titres against the target antigen. The peptide
CC immunogens of the invention are used to vaccinate against mammalian LHRH,
CC for use as (reversible) contraceptive; control of hormone-dependent
CC tumours (cancer of prostate or breast; also endometriosis); to prevent
CC boar taint (and improve meat quality) and for immunocastration. The
CC present sequence appears in the specification.

XX Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.5e-08;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPFKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 1 cgetyqsrvtphlpralmrstkc 25

RESULT 7
AAY80077
ID AAY80077 standard; peptide: 25 AA.

AC AAY80077;

DT 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Equus caballus.
OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -

XX

Query Match	46.9%	Score 68;	DB 18;	Length 20;
Best Local Similarity	76.5%	Pred. No. 0.0021;		
Matches 13;	Conservative	0;	Mismatches 4;	Indels 0;
QY	8	RVDHPHPFKPIVRSITK	24	
Db	3	rvthplpkdivrsiak	19	
RESULT	9			
AAV50893				
ID	AAV50893	standard; peptide;	17	AA.
XX				
XX	AAV50893;			
XX				
DT	24-FEB-2000	(first entry)		
XX				
DE	Antibody 15A.2	binding peptide 10	from Phbc7c	phage display library.
XX	Canine;	allergy; antibody 15A.2;	IgE; B cell;	mast cell; anti-allergy;
KW	Enzyme; prochylaxis;	treatment; mimotone		
KW				

OS
XX
XX
PN
XX

Synthetic.
EP957111-A2.

PD	17-NOV-1999.	
XX		
XX		
PF	09-APR-1999;	99EP-0107035.
XX		
XX		

PR 30-MAR-1999; 9905-0281/60.
XX
XX
PA (IDEX-) IDEXX LAB INC.
XX
XX

XX
XX
DR
XX
XX
PT
XX

WPI; 2000-040833/04.

Binding proteins used for treatment or prophylaxis of canine allergy -

PS disclosure: Flg 6; Jupp; English.
xx
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to native canine free or B-cell bound IgG.

to IgE when the IgE is bound to mast cells. The antibodies bind to defined epitopes on free or B-cell bound IgE molecules which have an important role in allergic reaction. The specific binding proteins are used to

CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.

xx	Sequence	17 AA;
xx	SQ	

Query Match	43.4%	Score b3:	DB 21;	Length 17;
Best Local Similarity	80.0%	Pred. No.	0.0097;	
Matches 12:	Conservative	0;	Mismatches 3;	Indels 0;
				Gaps 0;

QY	8	KVDHPHFPRPIVRSI	22
Db	3	rvthphlpkdivrsi	17

```
RESULT 10
AAAY50894
ID AAAY50894 standard: nentige: 17 AA
```

XX
XX
AC
XX
DT
AAY50894;
24-FEB-2000 (first entry)

XX	Antibody 15A.2 canine IgE binding epitope 1.
DE	Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
XX	epitope; prophylaxis; treatment; mimotope.
KW	Synthetic.
XX	
OS	
XX	
PN	EP957111-A2.
XX	17-NOV-1999.
PD	
PF	09-APR-1999; 99EP-0107035.
XX	
PR	09-APR-1998; 98US-0058331.
PR	30-MAR-1999; 99US-0281760.
XX	(IDEX-) IDEXX LAB INC.
PA	Lawton R, Mermer B, Francoeur G;
PI	WIPI; 2000-040833/04.
XX	
DR	Binding proteins used for treatment or prophylaxis of canine allergy -
PT	Disclosure; Fig 7; 30pp; English.
PS	
XX	This invention describes a novel binding protein which specifically
CC	binds to native canine free or B-cell bound IgE, and which doesn't bind
CC	to IgE when the IgE is bound to mast cells. The peptide products of the
CC	invention have anti-allergic activity. The antibodies bind to defined
CC	epitopes on free or B-cell bound IgE molecules which have an important
CC	role in allergic reaction. The specific binding proteins are used to
CC	produce a pharmaceutical composition, preferably with a diluent, which
CC	can be used for prophylaxis or treatment of canine allergy.
CC	AAI50876-fY5090 represent peptide mimotopes used in the method of the
CC	invention.
XX	
SQ	Sequence 17 AA;
Query Match	43.4%; Score 63; DB 21; Length 17;
Best Local Similarity	80.0%; Pred. No. 0.0097;
Matches 12; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Oy	8 RVDPHPFKPKIVRSI 22
Db	3 rvthphlpkdivrsi 17
RESULT 11	
AAV42585	ID AAV42585 standard; peptide; 22 AA.
ID	
XX	
AC	AAV42585;
XX	
DT	10-JAN-2000 (first entry)
DE	IgE peptide antagonist.
XX	
KW	Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
KW	receptor-binding; binding determinant sequence; anti-IgE antibody;
KW	allergic disease.
XX	Homo sapiens.
OS	
XX	US5965709-A.
PN	
PD	12-OCT-1999.
XX	
PF	21-APR-1994; 94US-0232539.
XX	
PR	14-AUG-1991; 91US-0744768.
PR	07-JAN-1994; 94US-0178583.
XX	(GETH) GENENTECH INC.
PA	Jardieu PM, Presta LG;
XX	WIPI; 1999-579941/49.
XX	Immunoglobulin E variants as peptide antagonists useful for raising and
PT	screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
PT	purification of FcepsilonRI receptor and in the treatment of allergic


```

AC  AAY50898;
XX
DT  24-FEB-2000 (first entry)
XX
DE  Antibody 15A.2 swine IgE binding epitope 1.
XX
KW  Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW  epitope; prophylaxis; treatment; mimotope.
XX
OS  Synthetic.
XX
PN  EP957111-A2.
XX
PD  17-NOV-1999.
XX
PF  09-APR-1999; 99EP-0107035.
XX
PR  09-APR-1998; 98US-0058331.
PR  30-MAR-1999; 99US-0281760.
XX
PA  (IDEX-) IDEXX LAB INC.
XX
PI  Lawton R, Mermer B, Francoeur G;
XX
DR  WPI; 2000-040833/04.
XX
PT  Binding proteins used for treatment or prophylaxis of canine allergy -
PS  Disclosure; Fig 7; 30pp; English.
XX
CC  This invention describes a novel binding protein which specifically
CC  binds to native canine free or B-cell bound IgE, and which doesn't bind
CC  to IgE when the IgE is bound to mast cells. The peptide products of the
CC  invention have anti-allergic activity. The antibodies bind to defined
CC  epitopes on free or B-cell bound IgE molecules which have an important
CC  role in allergic reaction. The specific binding proteins are used to
CC  produce a pharmaceutical composition. The method of invention, which
CC  can be used for prophylaxis or treatment of canine allergy.
CC  AAY50876-Y50900 represent peptide mimotopes used in the method of the
XX  invention.
SQ  Sequence 17 AA;

Query Match      37.2%; Score 54; DB 21; Length 17;
Best Local Similarity 71.4%; Pred. NO. 0.22;
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  9 VDPHPFPKPIVRSI 22
DB  4 vthpdlpkpilsr 17

Search completed: August 26, 2002, 10:18:42
Job time: 698 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:19:36 ; Search time 33.13 seconds
(without alignments)
18.432 Million cell updates/sec

Title: US-09-701-623c-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 134563

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	103	71.0	25	3	US-09-100-414B-95
2	103	71.0	25	4	US-09-303-323-55
3	62	42.8	22	2	US-08-232-539D-19
4	62	42.8	24	2	US-08-232-539D-20
5	37	25.5	15	2	US-08-630-645-7
6	37	25.5	15	5	PCT-US96-10220-7
7	35	24.1	13	6	5177197-5
8	35	24.1	20	2	US-08-564-972-54
9	35	24.1	20	2	US-08-564-972-55
10	34	23.4	13	4	US-09-258-754-218
11	34	23.4	13	4	US-09-042-107-218
12	34	23.4	18	4	US-08-602-999A-371
13	34	23.4	20	1	US-07-678-974D-5
14	34	23.4	20	2	US-08-945-168-10
15	33	22.8	17	4	US-08-602-999A-398
16	33	22.8	17	4	US-08-602-999A-399
17	33	22.1	18	4	US-09-177-249-298
18	32	22.1	18	4	US-09-461-697-30
19	32	22.1	18	5	PCT-US93-03748-8
20	32	22.1	20	1	US-07-678-974D-7
21	32	22.1	20	2	US-08-945-168-12
22	32	22.1	20	4	US-08-602-999A-120
23	32	22.1	20	4	US-08-630-915A-180
24	32	22.1	20	4	US-08-630-915A-222
25	31	21.4	11	2	US-08-630-645-8
26	31	21.4	11	5	PCT-US96-10220-8
27	31	21.4	16	1	US-08-447-010-14

28	31	21.4	18	1	US-08-395-602A-1	Sequence 1, Appli
29	31	21.4	18	2	US-08-021-625D-1	Sequence 1, Appli
30	31	21.4	19	1	US-08-116-733-12	Sequence 12, Appli
31	31	21.4	19	1	US-08-469-615-9	Sequence 9, Appli
32	31	21.4	19	1	US-08-466-763-9	Sequence 9, Appli
33	31	21.4	19	2	US-08-411-142A-9	Sequence 9, Appli
34	31	21.4	20	4	US-08-986-659B-8	Sequence 8, Appli
35	31	21.4	23	1	US-08-395-602A-2	Sequence 2, Appli
36	31	21.4	23	2	US-08-021-625D-2	Sequence 2, Appli
37	31	21.4	23	4	US-08-986-659B-9	Sequence 9, Appli
38	31	21.4	23	4	US-08-986-659B-33	Sequence 33, Appli
39	30	20.7	9	4	US-08-630-915A-61	Sequence 61, Appli
40	30	20.7	10	1	US-07-801-812A-8	Sequence 8, Appli
41	30	20.7	10	1	US-08-487-568-8	Sequence 8, Appli
42	30	20.7	12	1	US-08-433-318A-86	Sequence 86, Appli
43	30	20.7	12	2	US-08-922-048-86	Sequence 86, Appli
44	30	20.7	12	2	US-08-323-686-15	Sequence 15, Appli
45	30	20.7	12	5	PCT-US95-11127-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-95

Query Match 71.0%; Score 103; DB 3; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | :|| |||
Db 1 CGEYQSRVTHPHLPALMRSTTKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHR PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 71.0%; Score 103; DB 4; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | :|| |||
Db 1 CGETYQSRVTHPLPALMRSTKC 25

RESULT 3
US-08-232-539D-19
; Sequence 19, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: IgE Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-19

Query Match 42.8%; Score 62; DB 2; Length 22;
Best Local Similarity 55.6%; Pred. No. 0.0031;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIV 19
|| || || ||| | :
Db 5 GETYQCRVTHPLPALM 22

RESULT 4
US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: IgE Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-232-539D-20

Query Match 42.8%; Score 62; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 0.0034;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIV 19
||| ||| ||| | : :
DB 7 GETYQCRVTHPLPRALM 24

RESULT 5
US-08-630-645-7
Sequence 7, Application US/08630645
Patent No. 5948763
GENERAL INFORMATION:
APPLICANT: SOTO-JARA, Claudio
APPLICANT: BAUMANN, Marc
APPLICANT: FRANGIONE, Blas
TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: THEREOF FOR TREATMENT OF DISORDERS OR DISEASES ASSOCIATED
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,645
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,326
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: SOTO-JARA-1
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-630-645-7

Query Match 25.5%; Score 37; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 SRVDHPHPKPI 18
||| ||| ||| ||
DB 1 SRGDLPPFPVPI 12

RESULT 6
PCT-US96-10220-7
Sequence 7, Application PC/TUS9610220
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PEPTIDES AND PHARMACEUTICAL COMPOSITIONS
TITLE OF INVENTION: THEREOF FOR TREATMENT OF DISORDERS OR DISEASES ASSOCIATED
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 400
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10220
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,326
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/630,645
FILING DATE: 10-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: SOTO-JARA-1 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US96-10220-7

Query Match 25.5%; Score 37; DB 5; Length 15;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 SRVDHPHPKPI 18
||| ||| ||| ||
DB 1 SRGDLPPFPVPI 12

RESULT 7
5177197-5
Patent No. 5177197
APPLICANT: KANZAKI, TETSUTO;OLOFSSON, ANDERS;MOREN, ANITA;
WERNSTEDT, CHRISTER;HELLMAN, ULF;MIYAZONO, KOHEI;CLAESSON-WELSH,
LENA;HELDIN, CARL-HENRIK
TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
NUMBER OF SEQUENCES: 53
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/487,343
FILING DATE: 27-FEB-1990
SEQ ID NO: 5:
LENGTH: 13
5177197-5

Query Match 24.1%; Score 35; DB 6; Length 13;
Best Local Similarity 83.3%; Pred. No. 25;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEGYQ 6
||:||||
Db 3 CGQGYQ 8

RESULT 8
US-08-564-972-54
; Sequence 54, Application US/08564972
; Patent No. 5843462
; GENERAL INFORMATION:
; APPLICANT: Conti-Fine, B. M.
; TITLE OF INVENTION: DIPHTHERIA TOXIM EPITOPES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,972
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.344US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
US-08-564-972-54

Query Match 24.1%; Score 35; DB 2; Length 20;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYQSRVDH 11
||| |||
Db 13 GYQKTVDH 20

RESULT 9
US-08-564-972-55
; Sequence 55, Application US/08564972
; Patent No. 5843462
; GENERAL INFORMATION:

; APPLICANT: Conti-Fine, B. M.
; TITLE OF INVENTION: DIPHTHERIA TOXIM EPITOPES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,972
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.344US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:

; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
US-08-564-972-55

Query Match 24.1%; Score 35; DB 2; Length 20;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYQSRVDH 11
||| |||
Db 3 GYQKTVDH 10

RESULT 10
US-09-258-754-218
; Sequence 218, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Dipeptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 218
; LENGTH: 13
; TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-218

Query Match 23.4%; Score 34; DB 4; Length 13;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGYQSR 8
|||
Db 1 CGPGYQAO 8

RESULT 11
US-09-042-107-218
Sequence 218, Application US/09042107
Patent No. 6232287
GENERAL INFORMATION:
APPLICANT: Ruoslahti, Erkki
APPLICANT: Pasqualini, Renata
TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
FILE OF INVENTION: Tissues
FILE REFERENCE: P-LJ 2892
CURRENT APPLICATION NUMBER: US/09/042,107
CURRENT FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 436
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 218
LENGTH: 13
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-218

Query Match 23.4%; Score 34; DB 4; Length 13;
Best Local Similarity 62.5%; Pred. No. 36;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGYQSR 8
|||
Db 1 CGPGYQAO 8

RESULT 12
US-08-602-999A-371
Sequence 371, Application US/08602999A
Patent No. 6184205
GENERAL INFORMATION:
APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: QUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 371:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-371

Query Match 23.4%; Score 34; DB 4; Length 18;
Best Local Similarity 71.4%; Pred. No. 51;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPHFPKP 17
|||||
Db 1 HPHAPPP 7

RESULT 13
US-07-678-974D-5
Sequence 5, Application US/07678974D
Patent No. 5629146
GENERAL INFORMATION:
APPLICANT: DILLNER, JOAKIM
TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: BERMAN & AISENBERG
STREET: 1730 RHODE ISLAND AVENUE, N.W.,
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036-3186
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/678,974D
FILING DATE: 25-JUN-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: AISENBERG, Irwin M.
REGISTRATION NUMBER: 19,007
REFERENCE/DOCKET NUMBER: SG19171
TELEPHONE: 202-293-1404
TELEFAX: 202-872-0493
TELEX: 440 069 AIS UI
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-07-678-974D-5

Query Match 23.4%; Score 34; DB 1; Length 20;
Best Local Similarity 63.6%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 9 VDPHPFP--KP 17
| | | | |
Db 10 VGHYPFPIKPP 20

RESULT 14

US-08-945-168-10
; Sequence 10, Application US/08945168
; Patent No. 5989548
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,168
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SE96/00533
; FILING DATE: 23-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9501512-9
; FILING DATE: 24-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-945-168-10

Query Match 23.4%; Score 34; DB 2; Length 20;
Best Local Similarity 63.6%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 9 VDPHPFP--KP 17
| | | | |
Db 10 VGHYPFPIKPP 20

RESULT 15

US-08-602-999A-398
; Sequence 398, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.

; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 398:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-398

Query Match 22.8%; Score 33; DB 4; Length 17;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHPFKP 17
| | | | |
Db 4 HPNFQKP 10

Search completed: August 26, 2002, 10:19:37
Job time: 393 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:03:14 ; Search time 34.18 Seconds
(without alignments)
17.865 Million cell updates/sec

Title: US-09-701-623C-5
Perfect score: 140
Sequence: 1 CGETYQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backf1les1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	25	3	US-09-100-414B-95
2	140	100.0	25	4	US-09-303-323-55
3	140	100.0	42	3	US-09-100-414B-98
4	140	100.0	42	3	US-09-100-414B-99
5	140	100.0	42	3	US-09-100-414B-100
6	140	100.0	42	4	US-09-303-323-58
7	140	100.0	42	4	US-09-303-323-59
8	140	100.0	42	4	US-09-303-323-100
9	140	100.0	45	3	US-09-100-414B-101
10	140	100.0	45	4	US-09-303-323-101
11	140	100.0	46	3	US-09-100-414B-96
12	140	100.0	46	3	US-09-100-414B-97
13	140	100.0	46	4	US-09-303-323-96
14	140	100.0	46	4	US-09-303-323-97
15	140	100.0	63	3	US-09-100-414B-102
16	140	100.0	63	4	US-09-303-323-102
17	117	83.6	106	2	US-08-232-539D-54
18	117	83.6	113	2	US-08-232-539D-56
19	106.5	76.1	119	2	US-08-464-025A-1
20	99.5	71.1	109	4	US-08-466-163B-1
21	93	66.4	22	2	US-08-232-539D-19
22	93	66.4	24	2	US-08-232-539D-20
23	93	66.4	56	2	US-08-232-539D-18
24	89	63.6	118	3	US-08-466-151-1
25	85	60.7	426	1	US-08-336-583-2
26	85	60.7	426	5	PCT-US95-13795-2
27	49.5	35.4	1155	1	US-08-094-948A-29

28	49.5	35.4	1155	5	PCT-US96-09319-29	Sequence 29, Appl
29	46.5	33.2	1234	2	US-08-317-310A-15	Sequence 13, Appl
30	46.5	33.2	1234	5	PCT-US95-13041-15	Sequence 15, Appl
31	46	32.9	50	1	US-08-247-475-37	Sequence 37, Appl
32	46	32.9	50	1	US-08-479-650-37	Sequence 37, Appl
33	46	32.9	50	1	US-08-191-866D-59	Sequence 59, Appl
34	46	32.9	50	1	US-08-674-169-37	Sequence 37, Appl
35	46	32.9	50	2	US-08-185-949B-59	Sequence 59, Appl
36	46	32.9	91	2	US-09-047-125-31	Sequence 31, Appl
37	46	32.9	91	3	US-07-736-335E-31	Sequence 31, Appl
38	46	32.9	155	1	US-08-150-203A-10	Sequence 10, Appl
39	46	32.9	155	1	US-08-454-730-10	Sequence 10, Appl
40	45	32.1	109	2	US-08-646-981-6	Sequence 6, Appl
41	45	32.1	1120	4	US-09-147-404-1	Sequence 1, Appl
42	44	31.4	338	3	US-08-890-719-12	Sequence 12, Appl
43	44	31.4	355	3	US-08-890-719-11	Sequence 11, Appl
44	44	31.4	355	3	US-08-890-719-13	Sequence 13, Appl
45	44	31.4	483	4	US-09-049-672A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 100.0%; Score 140; DB 3; Length 25;
Best Local Similarity 100.0%; Pred No. 1,le-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGETYQSRVTHPLPALMRSTTKC 25
DB 1 CGETYQSRVTHPLPALMRSTTKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 100.0%; Score 140; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.le-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
DB 1 CGETYQSRVTHPLPALMRSTTKC 25

RESULT 3
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

Query Match 100.0%; Score 140; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.le-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
DB 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 4
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 100.0%; Score 140; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.le-16;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 5

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; FILING DATE: 20-JUNE-1998
; APPLICATION NUMBER: US/09/100,414B
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 100.0%; Score 140; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.le-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 6

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 100.0%; Score 140; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.le-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 7

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:

```

;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 42 amino acids
;
; TYPE: amino acid
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: peptide
US-09-303-323-99

```

Query Match: 100.0%; Score 140; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels

QY 1 CGETYQSRVTHPHLPRALMRSTTKC 25
|||
Db 18 CGETYQSRVTHPHLPRALMRSTTKC 42

```

RESULT      8
US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; DATE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.

```

```
Query Match      100.0%; Score 140; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels
```

Qy 1 CGETYQSRVTHPHLPALMRSTTKC 25
 |||||
 Db 18 CGETYQSRVTHPHLPALMRSTTKC 42

RESULT 9
US-09-100-414B-101
; Sequence 101, Application US/09100414B

```

: Patent No. 6025468
: GENERAL INFORMATION:
: APPLICANT: Wang, Chang Yi
: TITLE OF INVENTION: NOVEL LHRH PEPTIDE
: TITLE OF INVENTION: IMMUNOGENS
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morgan & Finnegan, L.L.P.
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10154-0054
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC Windows
: SOFTWARE: Word 97
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/100.414B
: FILING DATE: 20-JUNE-1998
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Maria H. Lin
: REGISTRATION NUMBER: 29,323
: REFERENCE/DOCKET NUMBER: 1151-4157
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-758-4800
: TELEFAX: 212-751-6849
: INFORMATION FOR SEQ ID NO: 101:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 45 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-09-100-414B-101

```

Query Match 100.0%; Score 140; DB 3; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 25; Conservative 0; Mismatches 0; Indels

Qy 1 CGETYQSRVTHPHLPRALMRSTTKC 25
| | | | | | | | | | | | | |
Db 21 CGETYQSRVTHPHLPRALMRSTTKC 45

```

RESULT 10
US-9-303-323-101
; Sequence 101, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 100.0%; Score 140; DB 4; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||

DB 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 11
US-09-100-414B-96
Sequence 96, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-96

Query Match 100.0%; Score 140; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25

DB 22 CGETYQSRVTHPLPALMRSTTKC 46
|||||

RESULT 12
US-09-100-414B-97
Sequence 97, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 100.0%; Score 140; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||

DB 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 13
US-09-303-323-96
Sequence 96, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible

2.

THIS PAGE BLANK (USPTO)

GenCore version 4.3
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:13:04 ; Search time 33.13 Seconds
(without alignments)
18.432 Million cell updates/sec

Title: US-09-701-623C-5
Perfect score: 140

Sequence: 1 CGETYSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 134663

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/1aa/5A.CONB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.CONB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.CONB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.CONB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	25	3	US-09-100-414B-95
2	140	100.0	25	4	US-09-303-323-95
3	93	66.4	22	2	US-08-232-539D-19
4	93	66.4	24	2	US-08-232-539D-20
5	35	25.0	24	1	US-07-976-358-24
6	35	25.0	25	1	US-07-976-358-15
7	35	25.0	25	1	US-07-976-358-18
8	35	25.0	25	1	US-07-976-358-21
9	34	24.3	25	1	US-07-976-358-17
10	33	23.6	18	2	US-09-017-205-9
11	32.5	23.2	15	2	US-08-671-094B-8
12	32.5	23.2	16	1	US-08-077-797A-14
13	32.5	23.2	16	5	PCT-US94-01238-14
14	32.5	23.2	17	1	US-08-311-307B-9
15	32	22.9	24	1	US-07-976-358-23
16	32	22.9	25	1	US-07-976-358-15
17	31	22.1	17	1	US-08-218-025A-186
18	31	22.1	18	2	US-09-017-205-52
19	31	22.1	21	4	US-09-077-991-5
20	30	21.4	13	1	US-08-466-285-6
21	30	21.4	16	4	US-08-602-999A-226
22	30	21.4	18	2	US-08-480-190-73
23	30	21.4	18	2	US-08-488-379-73
24	30	21.4	18	5	PCT-US93-07545-79
25	30	21.4	24	3	US-08-592-500-27
26	30	21.4	24	3	US-08-592-500-35
27	30	21.4	24	3	US-08-195-006-27

Sequence 35, Appl
Sequence 27, Appl
Sequence 35, Appl
Sequence 82, Appl
Sequence 3, Appl
Sequence 313, Appl
Sequence 313, Appl
Sequence 313, Appl
Sequence 160, Appl
Sequence 160, Appl
Sequence 371, Appl
Sequence 7, Appl
Sequence 12, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 37, Appl
Sequence 34, Appl

28 30 21.4 24 3 US-08-195-006-35
29 30 21.4 24 5 PCT-US94-07644A-27
30 30 21.4 24 5 PCT-US94-07644A-35
31 29.5 21.1 14 2 US-08-726-306A-82
32 29.5 21.1 18 4 US-09-146-580-3
33 29.5 21.1 24 2 US-08-637-759B-313
34 29.5 21.1 24 3 US-08-871-355A-313
35 29.5 21.1 24 4 US-09-201-945-313
36 29 20.7 11 1 US-08-548-540-160
37 29 20.7 11 5 PCT-US96-09809-160
38 29 20.7 18 4 US-08-602-999A-371
39 29 20.7 18 5 PCT-US92-07813-7
40 29 20.7 19 1 US-08-116-733-12
41 29 20.7 19 1 US-08-469-615-9
42 29 20.7 19 1 US-08-466-763-9
43 29 20.7 19 2 US-08-411-142A-9
44 29 20.7 20 1 US-08-440-861-37
45 29 20.7 20 2 US-08-564-972-34

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 35, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-95

Query Match 100.0%; Score 140; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGETYSRVTHPLPALMRSTTKC 25

DB 1 CGETYSRVTHPLPALMRSTTKC 25

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/017,205
;; FILING DATE: 02-FEB-1998
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mitchard, Leonard C
;; REGISTRATION NUMBER: 29,009
;; REFERENCE/DOCKET NUMBER: 604-436
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)816-4000
;; TELEFAX: (703)816-4100
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 18 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide from HSV-2 glycoprotein G
;; FRAGMENT TYPE: internal
US-09-017-205-9

Query Match 23.6%; Score 33; DB 2; Length 18;
Best Local Similarity 53.8%; Pred. No. 35;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 TYOSRVTHPLRP 16
Db 6 TYAARVYRLTR 18

RESULT 11
US-08-671-094B-8
; Sequence 8, Application US/08671094B
; Patent No. 591232
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Anti-Inflammatory Polypeptide
; TITLE OF INVENTION: Antagonists of Human Interleukin-8
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,094B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63086FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:

;; NAME/KEY: Modified-site
;; LOCATION: 4..6
;; OTHER INFORMATION: /note= "The cysteine residues
;; OTHER INFORMATION: may be substituted with aminobutyric acid, homocysteine or
;; OTHER INFORMATION: diaminosuberic acid."
US-08-671-094B-8

Query Match 23.2%; Score 32.5; DB 2; Length 15;
Best Local Similarity 53.8%; Pred. No. 34;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 CGETYOSRVTHPH 13
Db 4 CKITY-SKPFHPH 15

RESULT 12
US-08-077-797A-14
; Sequence 14, Application US/08077797A
; Patent No. 5679548
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F.
; APPLICANT: Rosenblum, Jonathan
; APPLICANT: Lerner, Richard A.
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
; TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5679548th Torrey Pines Road, TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,797A
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SC1276P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-077-797A-14

Query Match 23.2%; Score 32.5; DB 1; Length 16;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 2 GETYOSRVTHPHLP 15
Db 1 GDTHRGHLRH-HLP 13

```
RESULT 13
PCT-US94-01238-14
; Sequence 14, Application PC/TUS9401238
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
; BINDING SITES AND COMPOSITIONS THEREOF
; NUMBER OF SEQUENCES: 65
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01238
; FILING DATE: 01-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,797
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
PCT-US94-01238-14

Query Match 23.2%; Score 32.5; DB 5; Length 16;
Best Local Similarity 42.9%; Pred. No. 36;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy 2 GETYQSRVTHPLP 15
I::: :| | |
Db 1 GDTHRGHLRH-HLP 13

RESULT 14
US-08-311-307B-9
; Sequence 9, Application US/08311307B
; Patent No. 5627156
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human
; INTERLEUKIN-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,307B
; FILING DATE: 23-SEP-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100

us-09-701-623c-5.closed.rai

; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues may
; be substituted with aminobutyric acid, homocysteine or
; diaminosuberic acid."
US-08-311-307B-9

Query Match 23.2%; Score 32.5; DB 1; Length 17;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CGETYQSRVTHPH 13
I::| | | |
Db 6 CIKTY-SKPFHPH 17

RESULT 15
US-07-976-358-23
; Sequence 23, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-976-358-23

Query Match 22.9%; Score 32; DB 1; Length 24;
Best Local Similarity 42.9%; Pred. No. 72;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
```

OY 5 YQSRVTHPLRAL 18
| | : | | | :
Db 7 YSQPIRFLPRV 20

Search completed: August 26, 2002, 10:19:36
Job time: 392 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:07:04 ; Search time 79.43 Seconds
(Without alignments)
14.960 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLPRALRSYTKC 25

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 253524

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	140	100.0	25	AA199998	Optimised IGE-CH3
2	140	100.0	25	AA191212	Modified human Ige
3	140	100.0	25	AA196602	Peptide sequence O
4	108	77.1	25	AA199999	Optimised IGE-CH3
5	103	73.6	25	AA190000	Optimised IGE-CH3
6	94	67.1	25	AA190077	Optimised IGE-CH3
7	93	66.4	22	AA192585	Ige peptide antago
8	93	66.4	24	AA192586	Ige peptide antago
9	84	60.0	17	AA190895	Antibody 15A.2 hum
10	80	57.1	17	AA190896	Antibody 15A.2 gre
11	68	48.6	19	AA196503	Human Ige C epsilo

12	68	48.6	19	22	AA196830	Peptide PTX derive
13	68	48.6 <td>19 <td>22 <td>AA196831 <td>Ige peptide #11.</td> </td></td></td>	19 <td>22 <td>AA196831 <td>Ige peptide #11.</td> </td></td>	22 <td>AA196831 <td>Ige peptide #11.</td> </td>	AA196831 <td>Ige peptide #11.</td>	Ige peptide #11.
14	66	47.1 <td>13</td> <td>21 <td>AA196832 <td>Human Ige C epsilo</td> </td></td>	13	21 <td>AA196832 <td>Human Ige C epsilo</td> </td>	AA196832 <td>Human Ige C epsilo</td>	Human Ige C epsilo
15	66	47.1 <td>13</td> <td>21 <td>AA196833 <td>Peptide PI3b deriv</td> </td></td>	13	21 <td>AA196833 <td>Peptide PI3b deriv</td> </td>	AA196833 <td>Peptide PI3b deriv</td>	Peptide PI3b deriv
16	64	45.7 <td>20</td> <td>18 <td>AA196834 <td>Canine immunoglobi</td> </td></td>	20	18 <td>AA196834 <td>Canine immunoglobi</td> </td>	AA196834 <td>Canine immunoglobi</td>	Canine immunoglobi
17	64	45.7 <td>25</td> <td>21 <td>AA196835 <td>Optimised Ige-CH3</td> </td></td>	25	21 <td>AA196835 <td>Optimised Ige-CH3</td> </td>	AA196835 <td>Optimised Ige-CH3</td>	Optimised Ige-CH3
18	64	45.7 <td>25</td> <td>21 <td>AA196836 <td>Ige derived target</td> </td></td>	25	21 <td>AA196836 <td>Ige derived target</td> </td>	AA196836 <td>Ige derived target</td>	Ige derived target
19	60	42.9 <td>17</td> <td>21 <td>AA196837 <td>Antibody 15A.2 bin</td> </td></td>	17	21 <td>AA196837 <td>Antibody 15A.2 bin</td> </td>	AA196837 <td>Antibody 15A.2 bin</td>	Antibody 15A.2 bin
20	60	42.9 <td>17</td> <td>21 <td>AA196838 <td>Antibody 15A.2 can</td> </td></td>	17	21 <td>AA196838 <td>Antibody 15A.2 can</td> </td>	AA196838 <td>Antibody 15A.2 can</td>	Antibody 15A.2 can
21	58.5	41.8 <td>12</td> <td>21 <td>AA196839 <td>Human Ige C epsilo</td> </td></td>	12	21 <td>AA196839 <td>Human Ige C epsilo</td> </td>	AA196839 <td>Human Ige C epsilo</td>	Human Ige C epsilo
22	58.5	41.8 <td>12</td> <td>21 <td>AA196840 <td>Peptide PI3a deriv</td> </td></td>	12	21 <td>AA196840 <td>Peptide PI3a deriv</td> </td>	AA196840 <td>Peptide PI3a deriv</td>	Peptide PI3a deriv
23	55	39.3 <td>14</td> <td>21 <td>AA196841 <td>Human Ige C epsilo</td> </td></td>	14	21 <td>AA196841 <td>Human Ige C epsilo</td> </td>	AA196841 <td>Human Ige C epsilo</td>	Human Ige C epsilo
24	55	39.3 <td>14</td> <td>21 <td>AA196842 <td>Peptide PI3c deriv</td> </td></td>	14	21 <td>AA196842 <td>Peptide PI3c deriv</td> </td>	AA196842 <td>Peptide PI3c deriv</td>	Peptide PI3c deriv
25	52	37.1 <td>12</td> <td>21 <td>AA196843 <td>Human Ige C epsilo</td> </td></td>	12	21 <td>AA196843 <td>Human Ige C epsilo</td> </td>	AA196843 <td>Human Ige C epsilo</td>	Human Ige C epsilo
26	52	37.1 <td>12</td> <td>21 <td>AA196844 <td>Peptide PI3 derive</td> </td></td>	12	21 <td>AA196844 <td>Peptide PI3 derive</td> </td>	AA196844 <td>Peptide PI3 derive</td>	Peptide PI3 derive
27	52	37.1 <td>12</td> <td>21 <td>AA196845 <td>Ige peptide #16.</td> </td></td>	12	21 <td>AA196845 <td>Ige peptide #16.</td> </td>	AA196845 <td>Ige peptide #16.</td>	Ige peptide #16.
28	50	35.7 <td>17</td> <td>21 <td>AA196846 <td>Antibody 15A.2 fel</td> </td></td>	17	21 <td>AA196846 <td>Antibody 15A.2 fel</td> </td>	AA196846 <td>Antibody 15A.2 fel</td>	Antibody 15A.2 fel
29	48	34.3 <td>17</td> <td>21 <td>AA196847 <td>Antibody 15A.2 swi</td> </td></td>	17	21 <td>AA196847 <td>Antibody 15A.2 swi</td> </td>	AA196847 <td>Antibody 15A.2 swi</td>	Antibody 15A.2 swi
30	41.5	29.6 <td>16</td> <td>21 <td>AA196848 <td>Human apoptosis re</td> </td></td>	16	21 <td>AA196848 <td>Human apoptosis re</td> </td>	AA196848 <td>Human apoptosis re</td>	Human apoptosis re
31	40	28.6 <td>14</td> <td>22 <td>AA196849 <td>Human protein frag</td> </td></td>	14	22 <td>AA196849 <td>Human protein frag</td> </td>	AA196849 <td>Human protein frag</td>	Human protein frag
32	39	27.9 <td>21</td> <td>22 <td>AA196850 <td>Peptide #6058 enco</td> </td></td>	21	22 <td>AA196850 <td>Peptide #6058 enco</td> </td>	AA196850 <td>Peptide #6058 enco</td>	Peptide #6058 enco
33	39	27.9 <td>21</td> <td>22 <td>AA196851 <td>Protein #5684 enco</td> </td></td>	21	22 <td>AA196851 <td>Protein #5684 enco</td> </td>	AA196851 <td>Protein #5684 enco</td>	Protein #5684 enco
34	39	27.9 <td>21</td> <td>22 <td>AA196852 <td>Human bone marrow</td> </td></td>	21	22 <td>AA196852 <td>Human bone marrow</td> </td>	AA196852 <td>Human bone marrow</td>	Human bone marrow
35	39	27.9 <td>21</td> <td>22 <td>AA196853 <td>Human bone marrow</td> </td></td>	21	22 <td>AA196853 <td>Human bone marrow</td> </td>	AA196853 <td>Human bone marrow</td>	Human bone marrow
36	39	27.9 <td>21</td> <td>22 <td>AA196854 <td>Peptide #5714 enco</td> </td></td>	21	22 <td>AA196854 <td>Peptide #5714 enco</td> </td>	AA196854 <td>Peptide #5714 enco</td>	Peptide #5714 enco
37	39	27.9 <td>21</td> <td>22 <td>AA196855 <td>Peptide #6039 enco</td> </td></td>	21	22 <td>AA196855 <td>Peptide #6039 enco</td> </td>	AA196855 <td>Peptide #6039 enco</td>	Peptide #6039 enco
38	39	27.9 <td>25</td> <td>13 <td>AA196856 <td>Peptide corresp. t</td> </td></td>	25	13 <td>AA196856 <td>Peptide corresp. t</td> </td>	AA196856 <td>Peptide corresp. t</td>	Peptide corresp. t
39	37	26.4 <td>9</td> <td>21 <td>AA196857 <td>Monoclonal antibod</td> </td></td>	9	21 <td>AA196857 <td>Monoclonal antibod</td> </td>	AA196857 <td>Monoclonal antibod</td>	Monoclonal antibod
40	37	26.4 <td>9</td> <td>21 <td>AA196858 <td>Antibody 15A.2 bin</td> </td></td>	9	21 <td>AA196858 <td>Antibody 15A.2 bin</td> </td>	AA196858 <td>Antibody 15A.2 bin</td>	Antibody 15A.2 bin
41	37	26.4 <td>14</td> <td>21 <td>AA196859 <td>Immunogenic peptid</td> </td></td>	14	21 <td>AA196859 <td>Immunogenic peptid</td> </td>	AA196859 <td>Immunogenic peptid</td>	Immunogenic peptid
42	37	26.4 <td>23</td> <td>22 <td>AA196860 <td>Novel human diagno</td> </td></td>	23	22 <td>AA196860 <td>Novel human diagno</td> </td>	AA196860 <td>Novel human diagno</td>	Novel human diagno
43	37	26.4 <td>25</td> <td>19 <td>AA196861 <td>Human HUPF-I mutan</td> </td></td>	25	19 <td>AA196861 <td>Human HUPF-I mutan</td> </td>	AA196861 <td>Human HUPF-I mutan</td>	Human HUPF-I mutan
44	36	25.7 <td>9</td> <td>21 <td>AA196862 <td>Monoclonal antibod</td> </td></td>	9	21 <td>AA196862 <td>Monoclonal antibod</td> </td>	AA196862 <td>Monoclonal antibod</td>	Monoclonal antibod
45	36	25.7 <td>9</td> <td>21 <td>AA196863 <td>Antibody 15A.2 bin</td> </td></td>	9	21 <td>AA196863 <td>Antibody 15A.2 bin</td> </td>	AA196863 <td>Antibody 15A.2 bin</td>	Antibody 15A.2 bin

ALIGNMENTS

RESULT 1

AA199998

ID AA199998 standard; Peptide; 25 AA.

XX

AC AA199998;

XX

DT 15-MAY-2000 (first entry)

XX

DE Optimised Ige-CH3 domain antigen peptide for human Ige.

XX

KW Immunoglobulin E; Ige; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO9967293-A1.

XX

PD 29-DEC-1999.

XX

PF 21-JUN-1999; 99WO-US13959.

XX

PR 20-JUN-1998; 98US-0100287.

XX

PA (UNBI-) UNITED BIOMEDICAL INC.

XX

PI Wang CY, Walfield AM;

XX

WPI; 2000-160578/14.

XX

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions for immunisation against allergy

```

XX PS Claim 1; Page 21; 155pp; English.
XX CC
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies
XX CC specific for a target effector site on the epsilon-heavy chain of IgE,
XX CC and so preventing triggering and activation of mast cells and basophils
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe
XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX CC acid sequences used in the exemplification of the present invention.
XX CC
XX CC Sequence 25 AA;
XX CC
XX CC Query Match 100.0%; Score 140; DB 21; Length 25;
XX CC Best Local Similarity 100.0%; Pred. No. 7.4e-15;
XX CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX CC
QY 1 CGETYQSRVTHPHLPALMRSTTKC 25
DB 1 cgetyqsrvtphlpalmrsttkc 25
XX CC
RESULT 2
AAY91212
XX ID AAY91212 standard; peptide; 25 AA.
XX AC AAY91212;
XX XX
XX DT 22-MAY-2000 (first entry)
XX DE Modified human IgE CH3 domain, SEQ ID NO:92.
XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;
XX KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
XX KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
XX KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
XX KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
XX KW cholesterol ester transport protein; anti-arteriosclerotic.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO9966957-A2.
XX XX
XX PD 29-DEC-1999.
XX XX
XX PF 21-JUN-1999; 99WO-US13975.
XX XX
XX PR 20-JUN-1998; 98US-0100412.
XX XX
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX XX
XX PI Wang CY;
XX XX
XX DR WPI; 2000-160564/14.
XX XX
XX PT New artificial T helper cell epitope and derived immunogens with target
XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX PT or human immune deficiency virus -
XX XX
XX PS Example 6; Page 40; 129pp; English.

```

```

XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
XX CC and immunogenic peptides comprising the Th epitopes of the invention
XX CC along with B cell epitopes. The Th epitopes and peptide immunogens
XX CC containing them, are used to induce a T helper cell response,
XX CC specifically against Plasmodium falciparum, cholesterol ester transport
XX CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
XX CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX CC peptide immunogens may be used for prevention and/or treatment of
XX CC infections (HIV, foot-and-mouth disease or malaria); for cancer
XX CC immunotherapy; for inhibition of the action of luteinising hormone
XX CC releasing hormone (LHRH) for contraception, treatment of hormone-
XX CC dependent cancer, prevention of boar taint in meat, and
XX CC immunocastration; for promoting the growth of animals; or for
XX CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX CC Th (functional in genetically diverse subjects) into an immunogen
XX CC improves capacity to induce a strong T helper cell-mediated immune
XX CC response, resulting in production of antibodies against a target
XX CC antigen. Th can replace carrier proteins and pathogen-derived T helper
XX CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
XX CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
XX CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
XX CC from hepatitis B virus (HBV) surface antigen, and sequences
XX CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
XX CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
XX CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
XX CC is the LHRH target antigenic peptide used in these LHRH antigenic
XX CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
XX CC peptides comprising somatostatin and a Th epitope. Somatostatin
XX CC immunogens may be used to promote growth in livestock. AAY91208 is a
XX CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
XX CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
XX CC infection of T cells. AAY90212 is a modified version of a human IgE
XX CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
XX CC antigenic peptides which may be used in the treatment of allergies.
XX CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
XX CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
XX CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
XX CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
XX CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
XX CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
XX CC CETP peptide and a Th epitope which may be used to prevent or treat
XX CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
XX CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
XX CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
XX CC epitope which may be used as a component in an anti-HIV-1 vaccine.
XX CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
XX CC protein epitope from Yersinia species, and hinge spacer peptide, both of
XX CC which may optionally be used in the antigenic peptides of the
XX CC invention.
XX CC
XX CC Sequence 25 AA;
XX CC
XX CC Query Match 100.0%; Score 140; DB 21; Length 25;
XX CC Best Local Similarity 100.0%; Pred. No. 7.4e-15;
XX CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX CC
QY 1 CGETYQSRVTHPHLPALMRSTTKC 25
DB 1 cgetyqsrvtphlpalmrsttkc 25
XX CC
RESULT 3
AAY68602
XX ID AAY68602 standard; peptide; 25 AA.
XX AC AAY68602;
XX XX
XX DT 05-MAY-2000 (first entry)
XX XX
XX DE Peptide sequence of the invention.

```



```

XX 20-JUN-1998; 98US-0100287.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY, Walfield AM;
XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Claim 1; Page 99; 155pp; English.
XX
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX
XX Sequence 25 AA;

Query Match 73.6%; Score 103; DB 21; Length 25;
Best Local Similarity 68.0%; Pred. No. 4.3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGETYQSRVTHPLPALMRSTTKC 25
Db 1 cgegysrvdhpfpkpvrsitkc 25

RESULT 6
AAY80077
ID AAY80077 standard; Peptide; 25 AA.
XX
XX AAY80077;
XX
XX 15-MAY-2000 (first entry)
XX
XX Optimised IgE-CH3 domain antigen peptide for horse IgE.
XX
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX antibody; allergy; allergic disease; immunisation; anti-allergic;
XX anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
XX Equus caballus.
XX Synthetic.
XX
XX WO9967293-A1.
XX
XX 29-DEC-1999.
XX
XX 21-JUN-1999; 99WO-US13959.
XX
XX 20-JUN-1998; 98US-0100287.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY, Walfield AM;

```

```

XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX
XX Claim 1; Page 146; 155pp; English.
XX
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX
XX Sequence 25 AA;

Query Match 67.1%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 1.1e-07;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGETYQSRVTHPLPALMRSTTKC 25
Db 1 cgeyktstvwshpdlprevvrsiakc 25

RESULT 7
AAY42585
ID AAY42585 standard; peptide; 22 AA.
XX
XX AAY42585;
XX
XX 10-JAN-2000 (first entry)
XX
XX IgE peptide antagonist.
XX
XX Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
XX receptor-binding; binding determinant sequence; anti-IgE antibody;
XX allergic disease.
XX
XX Homo sapiens.
XX
XX US965709-A.
XX
XX 12-OCT-1999.
XX
XX 21-APR-1994; 94US-0232539.
XX
XX 14-AUG-1991; 91US-0744768.
XX
XX 07-JAN-1994; 94US-0178583.
XX
XX (GETH ) GENENTECH INC.
XX
XX Jardieu PM, Presta LG;
XX
XX WPI; 1999-579941/49.
XX
XX Immunoglobulin E variants as peptide antagonists useful for raising and
XX screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
XX purification of FcepsilonRI receptor and in the treatment of allergic
XX diseases -

```



```

RESULT 10
AA50896
ID AA50896 standard; peptide; 17 AA.
XX
AC AA50896;
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 green monkey IgE binding epitope 1.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
XX EP957111-A2.
XX
PN 17-NOV-1999.
XX
PD 09-APR-1999; 99EP-0107035.
XX
PF 09-APR-1998; 98US-0058331.
XX
PR 30-MAR-1999; 99US-0281760.
XX
XX (IDEX-) IDEXX LAB INC.
XX
XX Lawton R, Mermer B, Francoeur G;
XX
XX WPI; 2000-040833/04.
XX
XX Binding proteins used for treatment or prophylaxis of canine allergy -
XX
XX Disclosure; Fig 7; 30pp; English.
XX
XX This invention describes a novel binding protein which specifically
XX binds to native canine free or B-cell bound IgE, and which doesn't bind
XX to IgE when the IgE is bound to mast cells. The peptide products of the
XX invention have anti-allergic activity. The antibodies bind to defined
XX epitopes on free or B-cell bound IgE molecules which have an important
XX role in allergic reaction. The specific binding proteins are used to
XX produce a pharmaceutical composition, preferably with a diluent, which
XX can be used for prophylaxis or treatment of canine allergy.
XX
XX AA50876-150900 represent peptide mimotopes used in the method of the
XX invention.
XX
XX Sequence 17 AA;

Query Match 57.1%; Score 80; DB 21; Length 17;
Best Local Similarity 88.2%; Pred. NO. 1e-05;
Matches 15; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSRVTHPLPRALMRST 22
DB 1 qcrvthplpralvirst 17

RESULT 11
AAB26503
ID AAB26503 standard; peptide; 19 AA.
XX
AC AAB26503;
XX
DT 11-JAN-2001 (first entry)
XX
DE Human IgE C epsilon3/4 domain epitope P7.
XX
KW IgE; C epsilon3; C epsilon4; histamine release inhibitor; vaccine;
KW antibody; epitope; mimotope; human.
XX
OS Homo sapiens.
XX
XX WO200050461-A1.
PN

```

```

XX 31-AUG-2000.
PD
XX
XX 22-FEB-2000; 2000WO-EP01456.
XX
XX 25-FEB-1999; 99GB-0004408.
XX 21-JUL-1999; 99GB-0017144.
XX 07-AUG-1999; 99GB-0018598.
XX 07-AUG-1999; 99GB-0018599.
XX 07-AUG-1999; 99GB-0018601.
XX 07-AUG-1999; 99GB-0018604.
XX 07-AUG-1999; 99GB-0018606.
XX 29-OCT-1999; 99GB-0025618.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX (PEPT-) PEPTIDE THERAPEUTICS LTD.
XX
XX Friede M, Mason S, Turnell WG, Van Mechelen MP;
XX Vinals De Bassols YC;
XX
XX WPI; 2000-572074/53.
XX
XX Peptides comprising surface exposed epitopes or mimotopes derived from
XX C-epsilon3 or C-epsilon4 domains of IgE, useful for preventing or
XX treating allergy -
XX
XX 3; Page 4; 76pp; English.
XX
XX The present invention relates epitopes and mimotopes of an isolated
XX surface exposed epitope of C epsilon3 or C epsilon4 domain of IgE. The
XX epitopes were identified by calculating the accessible surface of each
XX IgE residue. Mimotopes were designed to be similar to the epitopes.
XX The epitopes are useful in preparing medicaments for treating or
XX preventing allergies. The epitopes and mimotopes of the invention
XX induce anti-IgE antibodies which are capable of raising
XX non-anaphylactic antibodies and inhibiting histamine release. The
XX present sequence is an IgE C epsilon3/4 domain epitope.
XX
XX Sequence 19 AA;

Query Match 48.6%; Score 68; DB 21; Length 19;
Best Local Similarity 100.0%; Pred. NO. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHLPRALMRSTTK 24
DB 1 phlpralmrsttk 13

RESULT 12
AAU16830
ID AAU16830 standard; peptide; 19 AA.
XX
AC AAU16830;
XX
XX 07-NOV-2001 (first entry)
XX
XX Peptide P7X derived from Cepsilon3/4 region of human IgE.
XX
XX Human; linkage technology; conjugated compound; carrier vehicle;
XX epitope; Cepsilon2; Cepsilon3; Cepsilon4; Immunoglobulin E;
XX IgE mediated disease; antibody response.
XX
XX Homo sapiens.
XX
XX WO200145745-A2.
XX
XX 28-JUN-2001.
XX
XX 21-DEC-2000; 2000WO-GB04935.
XX
XX 21-DEC-1999; 99GB-0030233.
XX

```

PR 22-FEB-2000; 2000GB-0004096.
PR 22-AUG-2000; 2000GB-0020707.
PR 22-AUG-2000; 2000GB-0020708.
XX
XX
PA (ACAM-) ACAMIS RES LTD.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Flinn N, Johnson T;
XX
XX WPI; 2001-521967/57.
DR
XX
XX A linkage comprising an immunogenic conjugate useful treatment of IgE mediated diseases -
PT
PT
XX
XX Example 5; Page 25; 48pp; English.
XX
XX The present invention relates to linkage methodology for use in the conjugation of compounds (e.g. peptides) to carrier vehicles (e.g. macromolecules, polymers, dendrimers, proteins) to produce biological and immunological constructs. The invention provides a method for linking an epitope (e.g. a peptide) to a carrier (e.g. a protein) for use in a pharmaceutical composition or a vaccine. The invention describes peptides derived from or mimotopes of the Cepsilon2, Cepsilon3 or Cepsilon4 regions of human immunoglobulin E (IgE) which are used to produce conjugated compounds. The compounds or compositions of the invention are useful in the manufacture of a medicament for the treatment of IgE mediated diseases. The invention allows for controlled conjugation of a peptide epitope (antigen) to a protein so as to form an immunogenic conjugate which may be able to raise a protective antibody response in an animal or human patient. CC AAU16632-AAU16913 represent peptides derived from or mimotopes of the Cepsilon2/Cepsilon3/Cepsilon4 region of human IgE.
XX
XX Sequence 19 AA;
SQ

Query Match 48.6%; Score 68; DE 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PHLPRALMRSTTK 24
Db 1 phlpralmrsttk 13
|||||

RESULT 13
AAB51033
ID AAB51033 standard; Peptide; 19 AA.
XX
XX
AC AAB51033;
XX
XX
DT 21-MAR-2001 (first entry)
XX
XX IgE peptide #11.
XX
DE Vaccine; immunoglobulin E; IgE; anti-allergy.
KW
XX
XX Mammalia.
OS
XX
XX WO200074716-A2.
XX
XX
PD 14-DEC-2000.
XX
XX
PF 06-JUN-2000; 2000WO-EP05164.
XX
XX
PR 08-JUN-1999; 99GB-0013327.
XX
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Prieels J;
XX
XX WPI; 2001-091150/10.
DR
XX

PT New vaccine comprising allergy peptides linked by an inert carrier, useful for boosting an anti-allergy immune response in an individual susceptible to an allergic response -
PT
XX
XX PS Claim 5; Page 21; 26pp; English.
XX
XX
CC The present invention relates to a composition comprising allergy peptides linked by an inert carrier. The allergy peptides are derived from immunoglobulin E (IgE) or IgE receptor. The present peptide is one such peptide from IgE. The composition is useful as a vaccine or for manufacturing a medicament for the prophylaxis or treatment of allergy. CC In particular, for boosting an anti-allergy immune response in an individual susceptible to an allergic response.
XX
XX Sequence 19 AA;
SQ

Query Match 48.6%; Score 68; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00087;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 PHLPRALMRSTTK 24
Db 1 phlpralmrsttk 13
|||||

RESULT 14
AAB26518
ID AAB26518 standard; Peptide; 13 AA.
XX
XX AAB26518;
XX
DT 11-JAN-2001 (first entry)
XX
XX Human IgE C epsilon3/4 domain mimotope P13b.
XX
KW IgE; C epsilon3; C epsilon4; histamine release inhibitor; vaccine;
KW antibody; epitope; mimotope; human.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1 /note= "N-terminal acetyl"
FT
XX
XX WO200050461-A1.
XX
XX 31-AUG-2000.
XX
XX 22-FEB-2000; 2000WO-EP01456.
XX
XX 25-FEB-1999; 99GB-0004408.
PR 21-JUL-1999; 99GB-0017144.
PR 07-AUG-1999; 99GB-0018598.
PR 07-AUG-1999; 99GB-0018599.
PR 07-AUG-1999; 99GB-0018601.
PR 07-AUG-1999; 99GB-0018604.
PR 07-AUG-1999; 99GB-0018606.
PR 29-OCT-1999; 99GB-0025618.
XX
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA (PEPT-) PEPTIDE THERAPEUTICS LTD.
XX
XX Friede M, Mason S, Turnell WG, Van Mechelen MP;
PI Vinals De Bassols YC;
XX
XX WPI; 2000-572074/53.
DR
XX
XX Peptides comprising surface exposed epitopes or mimotopes derived from C-epsilon3 or C-epsilon4 domains of IgE, useful for preventing or treating allergy -
PT
PT

PS Disclosure; Page 8; 76pp; English.

XX The present invention relates epitopes and mimotopes of an isolated
CC surface exposed epitope of C epsilon3 or C epsilon4 domain of IgE. The
CC epitopes were identified by calculating the accessible surface of each
CC IgE residue. Mimotopes were designed to be similar to the epitopes.
CC The epitopes are useful in preparing medicaments for treating or
CC preventing allergies. The epitopes and mimotopes of the invention
CC induce anti-IgE antibodies which are capable of raising
CC non-anaphylactic antibodies and inhibiting histamine release. The
CC present sequence is an IgE C epsilon3/4 domain mimotope.

XX SQ Sequence 13 AA;

Query Match 47.1%; Score 66; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 THPHLPRALMRS 21
| | | | | | | | | |
Db 1 thphlpralmrs 12

RESULT 15

AAU16849
ID AAU16849 standard; Peptide; 13 AA.

AC AAU16849;

XX 07-NOV-2001 (first entry)

DE Peptide P13b derived as mimotope of Cepsilon3/4 region of human IgE.

KW Human; linkage technology; conjugated compound; carrier vehicle;

KW epitope; Cepsilon2; Cepsilon3; Cepsilon4; immunoglobulin E;

KW IgE mediated disease; antibody response.

XX Homo sapiens.

OS Synthetic.

XX WO200145745-A2.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-GB04935.

XX 21-DEC-1999; 99GB-0030233.

XX 22-FEB-2000; 2000GB-0004096.

XX 22-AUG-2000; 2000GB-0020707.

XX 22-AUG-2000; 2000GB-0020708.

XX (ACAM-) ACAMBIS RES LTD.

PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Flinn N, Johnson T;

XX WPI; 2001-521967/57.

XX A linkage comprising an immunogenic conjugate useful treatment of IgE
PT mediated diseases -

XX Example 5; Page 25; 48pp; English.

XX The present invention relates to linkage methodology for use in the
CC conjugation of compounds (e.g. peptides) to carrier vehicles
CC (e.g. macromolecules, polymers, dendrimers, proteins) to produce
CC biological and immunological constructs. The invention provides a
CC method for linking an epitope (e.g. a peptide) to a carrier (e.g. a
CC protein) for use in a pharmaceutical composition or a vaccine. The
CC invention describes peptides derived from or mimotopes of the
CC Cepsilon2, Cepsilon3 or Cepsilon4 regions of human immunoglobulin E
CC (IgE) which are used to produce conjugated compounds. The compounds or

CC compositions of the invention are useful in the manufacture of a
CC medicament for the treatment of IgE mediated diseases. The invention
CC allows for controlled conjugation of a peptide epitope (antigen) to a
CC protein so as to form an immunogenic conjugate which may be able to
CC raise a protective antibody response in an animal or human patient.
CC AAU16632-AAU16913 represent peptides derived from or mimotopes of
CC the Cepsilon2/Cepsilon3/Cepsilon4 region of human IgE.

XX SQ Sequence 13 AA;

Query Match 47.1%; Score 66; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 THPHLPRALMRS 21
| | | | | | | | | |
Db 1 thphlpralmrs 12

Search completed: August 26, 2002, 10:18:42
Job time: 698 sec

THIS PAGE BLANK (USPTO)


```

RESULT 3
A61093
N:Alternate names: California mussel (fragments)
A:Alternate names: adhesive polyphenolic protein
C:Species: Mytilus californianus (California mussel)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: A61093
R:Waite, J.H.
J. Comp. Physiol. B 156, 491-496, 1986
A:Title: Mussel glue from Mytilus californianus Conrad: a comparative study.
A:Reference number: A61093; MUID:86279063
A:Accession: A61093
A:Molecule type: protein
A:Residues: 1-20 <WAI>
A:Note: the amino terminal residue was tentatively identified as Ser
A:Note: 12-Ser was also found
C:Comment: This glue protein is a component of the adhesive plaque of the byssus, the fi
.
C:Keywords: hydroxyproline; tandem repeat
F:4,14/Modified site: 3-hydroxyproline (Pro) #status experimental
F:5,9,15,19/Modified site: 4-hydroxyproline (Pro) #status experimental
F:7,17/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental
F:13/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental

Query Match 22.1%; Score 31; DB 2; Length 20;
Best Local Similarity 44.4%; Pred. No. 2.9e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 TYQSRVTHP 12
II : : : I
DB 6 TYKPKITVP 14

RESULT 4
B44379
omega-conotoxin SVIA - cone shell (Conus striatus)
N:Alternate names: SNX-157
C:Species: Conus striatus (striated cone)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997
C:Accession: B44379
R:Ramilo, C.A.; Zafaralla, G.C.; Nadasdi, L.; Hammerland, L.G.; Yoshikami, D.; Gray, W.R.
Biochemistry 31, 9919-9926, 1992
A:Title: Novel alpha- and omega-conotoxins from Conus striatus venom.
A:Reference number: A44379; MUID:93003172
A:Accession: B44379
A:Molecule type: protein
A:Residues: 1-24 <RAM>
A:Cross-references: CAS:137706-81-1
A:Experimental source: venom
A:Note: sequence extracted from NCBI backbone (NCBIP:116001); structure confirmed by che
C:Comment: This omega-conotoxin blocks presynaptic calcium channels.
C:Superfamily: omega-conotoxin
C:Keywords: acetylcholine release inhibition; amidated carboxyl end; calcium channel inh
F:1-15,8-18,14-23/Disulfide bonds: #status predicted
F:24/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 20.0%; Score 28; DB 2; Length 24;
Best Local Similarity 40.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTH 10
II : : : I
DB 15 CGRCYRGKCT 24

RESULT 5
S29749*
serum albumin - dog
C:Species: Canis lupus familiaris (dog)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S29749
R:Dixon, J.W.; Sarkar, B.
J. Biol. Chem. 249, 5872-5877, 1974
A:Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide
A:Reference number: S29749; MUID:75011422
A:Accession: S29749
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <DIX>
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match 20.0%; Score 28; DB 2; Length 24;
Best Local Similarity 44.4%; Pred. No. 9.8e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 ETYQSRVTH 11
II : : : I
DB 1 EAYKSEIAH 9

RESULT 6
PA0054
protein QF200017 - fungus (Fusarium sporotrichioides) (fragment)
C:Species: Fusarium sporotrichioides
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C:Accession: PA0054
R:Chow, L.P.; Fukaya, N.; Sugliura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A:Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotr
A:Reference number: PA0051
A:Accession: PA0054
A:Molecule type: protein
A:Residues: 1-15 <CHO>

Query Match 19.3%; Score 27; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 8.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRV 9
III I : I
DB 3 GETQQTQKV 10

RESULT 7
S72535
probable acr-2 regulatory leader protein - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 21-Jul-2000
C:Accession: S72535
R:Akiyama, M.; Nakashima, H.
Biochim. Biophys. Acta 1307, 187-192, 1996
A:Title: Molecular cloning of the acr-2 gene which controls acriflavine sensitivity 1
A:Reference number: S72535; MUID:96283814
A:Accession: S72535
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-23 <AKI>
A:Cross-references: EMBL:D45893; NID:g1754593; PIDN:BAA08306.1; PTD:gl754594

Query Match 19.3%; Score 27; DB 2; Length 23;
Best Local Similarity 28.6%; Pred. No. 1.3e+03;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 12 PHLPRALMRSTTKC 25
I : I : I
DB 7 PYQPEMIRLEVSSC 20

RESULT 8
```

S32551
glutathione transferase (EC 2.5.1.18) mu (isoform pI 6.4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 05-Jan-1996
C:Accession: S32551; S32550
R:Singhal, S.S.; Saxena, M.; Ahmad, H.; Awasthi, Y.C.
Biochim. Biophys. Acta 1116, 137-146, 1992
A:Title: Glutathione S-transferases of mouse liver: sex-related differences in the expression
A:Reference number: S32548; MUID:92256466
A:Accession: S32551
A:Molecule type: protein
A:Residues: 1-13 <SIN1>
A:Experimental source: female
A:Accession: S32550
A:Molecule type: protein
A:Residues: 1-13 <SIN2>
A:Experimental source: male
C:Keywords: transferase

Query Match 18.6%; Score 26; DB 2; Length 13;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RVTTP 12
Db 9 RLTHP 13

RESULT 9
S62374
alpha-1-antichymotrypsin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 31-Mar-1997
C:Accession: S62374
R:Tsuda, M.; Sei, Y.; Ohkubo, T.; Yamamura, M.; Kamiguchi, H.; Akatsuka, A.; Tsuda, T.;
Eur. J. Biochem. 235, 821-827, 1996
A:Title: The defective secretion of a naturally occurring alpha-1-antichymotrypsin variant
A:Reference number: S62374; MUID:96184564
A:Accession: S62374
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-14 <TSU>

Query Match 18.6%; Score 26; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 YQSRVTHP 12
Db 4 FMSKVTNP 11

RESULT 10
A60621
somatotropin - Atlantic salmon (fragment)
N:Alternate names: growth hormone
C:Species: Salmo salar (Atlantic salmon)
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C:Accession: A60621
R:Skibelli, V.; Andersen, O.; Gautvik, K.M.
Gen. Comp. Endocrinol. 80, 333-344, 1990
A:Title: Purification and characterization of Atlantic salmon growth hormone and evidence
A:Reference number: A60621; MUID:91146880
A:Accession: A60621
A:Molecule type: protein
A:Residues: 1-25 <SKT>
A:Note: this protein displayed charge heterogeneity from variable levels of phosphorylation
C:Superfamily: prolactin
C:Keywords: hormone; phosphoprotein; pituitary

Query Match 18.6%; Score 26; DB 2; Length 25;
Best Local Similarity 62.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 SRVTHPHL 14
Db 12 NRVQHLHL 19

RESULT 11
I67422
gene GLUT4 protein - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 04-Sep-1998
C:Accession: I67422
R:Olson, A.L.; Edgington, N.P.; Moye-Rowley, W.S.; Pessin, J.E.
Endocrinology 136, 1962-1968, 1995
A:Title: Characterization of 5'-heterogeneity of the rat GLUT4/muscle-adipose glucose
A:Reference number: I53293; MUID:95237122
A:Accession: I67422
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-25 <RES>
A:Cross-references: GB:S77784; NID:g998977
C:Genetics:
A:Gene: GLUT4
C:Superfamily: glucose transport protein

Query Match 18.6%; Score 26; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.1e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GETYQSRVT 10
Db 13 GEPPQQRVT 21

RESULT 12
A58647
alphaA-conotoxin PIVA [validated] - cone shell (Conus purpurascens)
C:Species: Conus purpurascens (purple cone)
C:Date: 31-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 15-Sep-2000
C:Accession: A58647
R:Hopkins, C.; Grilley, M.; Miller, C.; Shon, K.J.; Cruz, L.J.; Gray, W.R.; Dykert, J.
J. Biol. Chem. 270, 22361-22367, 1995
A:Title: A new family of Conus peptides targeted to the nicotinic acetylcholine receptor
A:Reference number: A58647; MUID:95403432
A:Accession: A58647
A:Molecule type: protein
A:Residues: 1-25 <HOP>
R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.
submitted to the Brookhaven Protein Data Bank, December 1996
A:Reference number: A67666; PDB:1PIp
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, resi
R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.
Biochemistry 36, 1669-1677, 1997
A:Title: NMR structure determination of a novel conotoxin, [Pro 7.13] alpha A-conotoxin
A:Reference number: A58646; MUID:97200721
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
C:Superfamily: unassigned conotoxins
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline;
F:2-16,3-11,14-23/Disulfide bonds: #status experimental
F:7,13/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F:20/Modified site: 4-hydroxyproline (Pro) #status experimental
F:25/Modified site: amidated carboxyl end (Gln) #status experimental

Query Match 18.2%; Score 25.5; DB 2; Length 25;
Best Local Similarity 41.7%; Pred. No. 2.5e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 1 CGETYQSRVTHP 12

C:Genetics:
A:Gene: AGU

|| : | : ||
Db 3 CG-SYPNAACHP 13

RESULT 13

H37196
bradykinin-potentiating peptide 8 - island jararaca
C:Species: Bothrops insularis (island jararaca)
C:Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994
C:Accession: H37196
R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides f
A:Reference number: A37196; MUID:90351557
A:Accession: H37196
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <CIN>
C:Keywords: pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 17.9%; Score 25; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 11 HPPLP 15
|| : | : ||
Db 5 HPNIP 9

RESULT 14

A56901
nerve growth factor - puff adder (fragment)
C:Species: Bitis arietans (puff adder)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 20-Jun-2000
C:Accession: A56901
R:Smith, P.J.; Brandt, W.F.; Stickells, B.J.; von Holt, C.
Comp. Biochem. Physiol. B 103, 975-980, 1992
A:Title: Bitis arietans nerve growth factor is a disulphide-linked homodimer.
A:Reference number: A56901; MUID:93121763
A:Accession: A56901
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-21 <SMI>
A:Note: sequence extracted from NCBI backbone (NCBIP:122566)
C:Keywords: disulfide bond; homodimer; venom

Query Match 17.9%; Score 25; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 ETYQSRVTHP 12
| | | | |
Db 7 EYIQRVCQP 16

RESULT 15

I37144
aspartylglycosaminuria - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I37144
R:Park, H.; Vettese, M.B.; Fensom, A.H.; Fisher, K.J.; Aronson, N.N.
Biochem. J. 290, 735-741, 1993
A:Title: Characterization of three alleles causing aspartylglycosaminuria: two from a Br
A:Reference number: I37144; MUID:93207523
A:Accession: I37144
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-22 <RES>
A:Cross-references: EMBL:X73071; NID:g312227; PIDN:CAA51529.1; PID:g312228

Query Match 17.9%; Score 25; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 2.6e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TYQSRVTH 11
| : | : |
Db 5 TFSRRVSH 12

Search completed: August 26, 2002, 10:20:41
Job time: 347 sec

14-00000

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:19:40 ; Search time 22.95 Seconds
(without alignments)
42.178 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPHLPALMRSTTKC 25

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 1503

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	26	18.6	21	1	CPFA_TREPH
2	25.5	18.2	25	1	CXA4_CONPU
3	25	17.9	10	1	BPB8_BOVIN
4	25	17.9	20	1	MI17_BOVIN
5	24	17.1	10	1	BPB2_BOVIN
6	24	17.1	15	1	MCA2_RHOOP
7	24	17.1	19	1	H170_RAT
8	23	16.4	14	1	CXA1_CONMA
9	23	16.4	20	1	UN05_PINP;
10	23	16.4	21	1	FAV1_TETPY
11	23	16.4	23	1	SODM_RANCA
12	23	16.4	24	1	BRL_RANBP
13	23	16.4	24	1	GAP6_RANRU
14	22	15.7	10	1	MALE_KLEPN
15	22	15.7	16	1	YMRP_PSEPU
16	22	15.7	20	1	MIF_PIG
17	22	15.7	21	1	TERT_APIME
18	22	15.7	24	1	BRIA_RANES
19	21.5	15.4	22	1	TX12_TRIWA
20	21	15.0	13	1	NEUT_CAVPO
21	21	15.0	14	1	CXA1_CONCN
22	21	15.0	17	1	MAOX_CHICK
23	21	15.0	17	1	APID_BOMPA
24	21	15.0	18	1	PHI_PERAM
25	21	15.0	19	1	CXA2_CONST
26	21	15.0	19	1	FRHG_METBA
27	21	15.0	20	1	CS21_STRTR
28	21	15.0	20	1	TL18_SPIOL
29	21	15.0	21	1	THAN_PODMA
30	21	15.0	24	1	FEDG_AMEYE
31	21	15.0	24	1	Y3KD_NEUCR
32	20	14.3	12	1	UR2_POLSP
33	20	14.3	13	1	CXA1_CONST

34	20	14.3	15	1	ACEA_ACICA	P28467 acinetobact
35	20	14.3	15	1	CXA1_CONGE	P01519 conus geogr
36	20	14.3	15	1	SODM_ENTAE	P22799 enterobacte
37	20	14.3	15	1	UN01_PINPS	P81106 pinus pinas
38	20	14.3	16	1	CXA1_CONEP	P56638 conus episc
39	20	14.3	18	1	A2M_OCTVU	P30800 octopus vul
40	20	14.3	20	1	SODF_PASPI	P81527 pasteurella
41	20	14.3	23	1	ALL5_HORSE	P82616 equus cabal
42	20	14.3	24	1	LPER_STRFR	P45440 streptomyce
43	20	14.3	25	1	GBX1_MOUSE	P82976 mus musculu
44	20	14.3	25	1	RL41_SCHPO	Q94710 schizosacch
45	19	13.6	9	1	LMT3_LOCHI	P41489 locusta mig

ALIGNMENTS

RESULT 1
ID CPFA_TREPH STANDARD; PRT; 21 AA.
AC P56738;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Cytoplasmic filament protein A (fragment).
GN CPFA.
OS Treponema phagedenis.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=162;
RN [1]
RP SEQUENCE.
RC STRAIN-KAZAN 5;
RX MEDLINE=96236033; PubMed=8655496;
RA You Y., Elmore S., Colton L.L., Mackenzie C., Stoops J.K.,
RA Weinstock G.M., Norris S.J.;
RT "Characterization of the cytoplasmic filament protein gene (cfpa) of
Treponema pallidum subsp. pallidum.";
RL J. Bacteriol. 178:3177-3187(1996).
CC -!- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
CC -!- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
APPPOSITION TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
KW Structural protein; Antigen.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2231 MW; 574604B4FFC2D017 CRC64;

Query Match 18.6%; Score 26; DB 1; Length 21;
Best Local Similarity 50.0%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 9 VTHPHLPAL 18
| | | | |
Db 11 VFHPEKPSAV 20

RESULT 2
ID CXA4_CONPU STANDARD; PRT; 25 AA.
AC P55963;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-A conotoxin PIVA.
OS Conus purpurascens (Purple cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OX Neogastropoda; Conoidea; Conidae; Conus.
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=95403432; PubMed=7673220;

Page 4

RC TISSUE=Skin;
RX MEDLINE=93080555; PubMed=1449472;
RA Morikawa N., Hagiwara K., Nakajima T.;
RT "Brevinin-1 and -2, unique antimicrobial peptides from the skin of
the frog, Rana brevipoda porsa.";
RL Biochem. Biophys. Res. Commun. 189:184-190(1992).
CC !- FUNCTION: SHOWS ANTIBACTERIAL ACTIVITY AGAINST REPRESENTATIVE
GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIAL SPECIES, AND A VERY
HIGH HEMOLYTIC ACTIVITY.
CC !- SUBCELLULAR LOCATION: Secreted.
CC !- TISSUE SPECIFICITY: SKIN.
CC !- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
FAMILY.
DR PIR: JCI1355;
KW Amphibian skin; Antibiotic; Hemolysis.
FT DISULFID 18 24 BY SIMILARITY.
SQ SEQUENCE 24 AA: 2531 MW; C866285B191EFD4 CRC64;

Query Match 16.4%; Score 23; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 1.8e+03;
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 14 LPRALMRSTTKC 25
: : : : :
Db 13 VPALFCKITKCC 24

RESULT 13
GAE6_RANRU STANDARD; PRT; 24 AA.
AC P80400;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gaegurin-6.
OS Rana rugosa (Wrinkled frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Ranidae; Rana.
OX NCBI_TaxID=8410;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=95091844; PubMed=7999137;
RA Park J.M., Jung J.-E., Lee B.J.;
RT "Antimicrobial peptides from the skin of a Korean frog, Rana
rugosa.";
RL Biochem. Biophys. Res. Commun. 205:948-954(1994).
CC !- FUNCTION: HAS A NON-HEMOLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF
ACTIVITY AGAINST BOTH GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA,
FUNGI AND PROTOZOA.
CC !- SUBUNIT: MONOMER.
CC !- SUBCELLULAR LOCATION: Secreted.
CC !- TISSUE SPECIFICITY: SKIN.
CC !- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
FAMILY.
KW Amphibian skin; Antibiotic.
FT DISULFID 18 24 BY SIMILARITY.
SQ SEQUENCE 24 AA: 2610 MW; 09918123FF90CCFD CRC64;

Query Match 16.4%; Score 23; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 1.8e+03;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 14 LPRALMRSTTKC 25
: : : : :
Db 13 LPTICKISYKC 24

RESULT 14
MALE_KLEPN STANDARD; PRT; 10 AA.
ID MALE_KLEPN

AC Q05564;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Maltose-binding periplasmic protein (Maltodextrin-binding protein)
(MBBP) (Fragment).
GN MALE.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1033-5P14 / KAY2026;
RX MEDLINE=93211295; PubMed=8459773;
RA Bachelier S., Perrin D., Hofnung M., Gilson E.;
RT "Bacterial interspersed mosaic elements (BIMes) are present in the
genome of Klebsiella.";
RL Mol. Microbiol. 7:537-544(1993).
CC !- FUNCTION: MALE IS INVOLVED IN THE HIGH-AFFINITY MALTOSE MEMBRANE
TRANSPORT SYSTEM. INITIAL RECEPTOR FOR THE ACTIVE TRANSPORT OF AND
CHEMOTAXIS TOWARD MALTOLOGOSACCHARIDES.
CC !- SUBCELLULAR LOCATION: Periplasmic.
CC !- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
PROTEIN FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X68329; CAA48406.1; -.
DR InterPro; IPR000567; SBP_bac.1.
DR PROSITE; PS01037; SBP_BACTERIAL.1; PARTIAL.
KW Transport; Sugar transport; Periplasmic.
FT NON_TER 1
SQ SEQUENCE 10 AA: 1159 MW; 8FD8DC4415A6DDDA CRC64;

Query Match 15.7%; Score 22; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. No. 1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 QSRVT 10
: : : : :
Db 5 QSRIT 9

RESULT 15
YMOR_PSEPU STANDARD; PRT; 16 AA.
AC Q02210;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in MORA 3' region (Fragment).
OS Pseudomonas putida.
OG Plasmid pMDH7.2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M10;
RX MEDLINE=93199531; PubMed=8452544;
RA Willey D.L., Caswell D.A., Lowe C.R., Bruce N.C.;
RT "Nucleotide sequence and over-expression of morphine dehydrogenase, a
plasmid-encoded gene from Pseudomonas putida M10.";
RL Biochem. J. 290:539-544(1993).
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M94775; AAB17357.1; -
 DR PIR; S30384; S30384.
 KW Hypothetical protein; plasmid.
 FT NON_TER 16
 SQ SEQUENCE 16 AA; 1962 MW; A2F1EB8C172766ED CRC64;

Query Match 15.7%; Score 22; DB 1; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.7e+03;
 Matches 6; Conservative 1; Mismatches 3; Indels 2; Gaps 1;
 QY 14 LPRALMRSTTKC 25
 |||: |||
 Db 6 LPRS--RSLRSC 15

Search completed: August 26, 2002, 10:30:17
 Job time: 637 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	31	22.1	17	2	Q34216	Q34216	sphingomona
2	38	20.0	17	6	Q9TR98	Q9TR98	canis famil
3	28	20.0	23	2	P95839	P95839	staphylococ
4	27	19.3	22	13	Q910C6	Q910C6	gallinula c
5	27	19.3	22	13	Q902V4	Q902V4	rallus limi
6	27	19.3	22	13	Q902V3	Q902V3	rallus long
7	27	19.3	22	13	Q902V2	Q902V2	rallus long
8	27	19.3	25	4	Q9UBT0	Q9UBT0	homo sapien
9	26	18.6	18	4	Q9BQT0	Q9BQT0	homo sapien
10	26	18.6	18	4	Q13665	Q13665	homo sapien
11	26	18.6	20	6	Q9TQ26	Q9TQ26	macaca fasc
12	26	18.6	23	2	Q52321	Q52321	escherichia
13	26	18.6	24	15	Q73337	Q73337	human immun
14	25.5	18.2	19	13	Q42416	Q42416	gallus gall
15	25.5	18.2	22	4	Q9BUH0	Q9BUH0	homo sapien
16	25	17.9	17	4	Q16310	Q16310	homo sapien

DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ALBUMIN (FRAGMENT)
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE
RX MEDLINE=95007849; PubMed=7923441;
RA Miller M.J., Parmelee D.C., Benjamin T., Sechi S., Dooley K.L.,
RA Kadlubar F.F.;
RT "Plasma proteins as early biomarkers of exposure to carcinogenic
aromatic amines";
RL Chem.-Biol. Interact. 93:221-234(1994).
SQ SEQUENCE 17 AA; 2024 MW; 1D39F70F7D23B269 CRC64;

Query Match 20.0%; Score 28; DB 6; Length 17;
Best Local Similarity 44.4%; Pred. No. 9e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 ETYQSRVTH 11
| | | : |
Db 1 EAYKSEIAH 9

RESULT 3
P95839 ID P95839 PRELIMINARY; PRT; 23 AA.
AC P95839;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
DE ORF56 (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=98427339; PubMed=8830703;
RA Wu S., de Lencastre H., Tomasz A.;
RT "Sigma-B, a putative operon encoding alternate sigma factor of
Staphylococcus aureus RNA polymerase: molecular cloning and DNA
sequencing";
RL J. Bacteriol. 178:6036-6042(1996).
DR EMBL; Y09929; CAA71063.1; -;
FT NON_TER 1 1
SQ SEQUENCE 23 AA; 2541 MW; 7F47717B1767D34F CRC64;

Query Match 20.0%; Score 28; DB 2; Length 23;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGETYS 7
| | | | |
Db 11 CNETYLS 17

RESULT 4
Q910C6 ID Q910C6 PRELIMINARY; PRT; 22 AA.
AC Q910C6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ADENYLATE KINASE (FRAGMENT).
OS Gallinula chloropus (Common moorhen).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Gallinula.
OX NCBI_TaxID=9123;
RN [1]
RP SEQUENCE FROM N.A.
RA Shapero L.H., Dumbacher J.P.;
RT "Adenylate kinase intron 5: a new nuclear locus for avian
systematics";
RL Auk 118:248-255(2001).
DR EMBL; AF372551; AAK43728.1; -;
KW Kinase.
FT NON_TER 1 1
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2447 MW; CB0CFBB55B82C2A8 CRC64;

Query Match 19.3%; Score 27; DB 13; Length 22;
Best Local Similarity 30.8%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHL 14
| | : : : |
Db 2 GEEFEKKIAPPTL 14

RESULT 5
Q90ZV4 ID Q90ZV4 PRELIMINARY; PRT; 22 AA.
AC Q90ZV4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ADENYLATE KINASE (FRAGMENT).
OS Rallus limicola.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Rallus.
OX NCBI_TaxID=156759;
RN [1]
RP SEQUENCE FROM N.A.
RA Shapero L.H., Dumbacher J.P.;
RT "Adenylate kinase intron 5: a new nuclear locus for avian
systematics";
RL Auk 118:248-255(2001).
DR EMBL; AF307900; AAK43538.1; -;
KW Kinase.
FT NON_TER 1 1
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2447 MW; CB0CFBB55B82C2A8 CRC64;

Query Match 19.3%; Score 27; DB 13; Length 22;
Best Local Similarity 30.8%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHL 14
| | : : : |
Db 2 GEEFEKKIAPPTL 14

RESULT 6
Q90ZV3 ID Q90ZV3 PRELIMINARY; PRT; 22 AA.
AC Q90ZV3;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ADENYLATE KINASE (FRAGMENT).
OS Rallus longirostris (Clapper rail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Rallus.
OX NCBI_TaxID=54361;
RN [1]
RP SEQUENCE FROM N.A.

RA Shapiro L.H., Dumbacher J.P.;
RT "Adenylate kinase intron 5: a new nuclear locus for avian
RT systematics";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF307901; AAK43539.1; -
KW Kinase.
FT NON_TER 1 1
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2447 MW; CB0CFB5B5B82C2A8 CRC64;

Query Match 19.3%; Score 27; DB 13; Length 22;
Best Local Similarity 30.8%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHL 14
|| : : | |
Db 2 GEEFEKKIAPPTL 14

RESULT 7
Q90ZV2 ID Q90ZV2 PRELIMINARY; PRT; 22 AA.
AC Q90ZV2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ADENYLATE KINASE (FRAGMENT).
OS Sarothrura elegans.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Gruiformes; Rallidae; Sarothrura.
OX NCBI_TaxID=156766;
RN [1]
RP SEQUENCE FROM N.A.
RA Shapiro L.H., Dumbacher J.P.;
RT "Adenylate kinase intron 5: a new nuclear locus for avian
RT systematics";
RL Auk 118:248-255(2001).
DR EMBL; AF307902; AAK43540.1; -
KW Kinase.
FT NON_TER 1 1
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2444 MW; 7F0CFB5B5B82C2B5 CRC64;

Query Match 19.3%; Score 27; DB 13; Length 22;
Best Local Similarity 30.8%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHL 14
|| : : | |
Db 2 GEEFEKKIAPPTL 14

RESULT 8
Q9UBT0 ID Q9UBT0 PRELIMINARY; PRT; 25 AA.
AC Q9UBT0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE 6-PHOSPHOFRUCTO-2-KINASE/FRUCTOSE-2, 6-BISPHOSPHATASE (EC 2.7.1.105)
DE (FRAGMENT).
GN PFK2 OR IPFK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fukasawa M., Sakakibara R., Yamamoto N.;
RC TISSUE-PLACENTA;
RT "Identification of the Promoter Region of Human Placental 6-

RT Phosphofructo-2-kinase/Fructose-2,6-bisphosphatase Gene.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Mahlnecht U., Bucala R.;
RT "Homo sapiens inducible 6-phosphofructo-2-kinase/fructose-2,6-
RT bisphosphatase (IPFK2), genomic sequence";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB033994; BAA89351.1; -
DR EMBL; AF110958; AAD23989.1; -
KW Kinase; Transferase.
FT NON_TER 25 25
SQ SEQUENCE 25 AA; 2984 MW; 4ACC24B43FDBB74C CRC64;

Query Match 19.3%; Score 27; DB 4; Length 25;
Best Local Similarity 40.9%; Pred. No. 2e+03;
Matches 9; Conservative 0; Mismatches 5; Indels 8; Gaps 1;

QY 3 ETYQSRVT-----HPHLPR 16
| | | | |
Db 4 ELTQSRVQKIWVPVDRPFLPR 25

RESULT 9
Q9BQT0 ID Q9BQT0 PRELIMINARY; PRT; 18 AA.
AC Q9BQT0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 2.2 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21195339; PubMed=11297743;
RA Holzmann K., Ambrosch I., Elbling L., Micksche M., Berger W.;
RT "A small upstream open reading frame causes inhibition of human major
RT vault protein expression from a ubiquitous mRNA splice variant.";
RL FEBS Lett. 494:99-104(2001).
DR EMBL; AJ291367; CAC35315.1; -
KW Hypothetical protein.
SQ SEQUENCE 18 AA; 2179 MW; 5D06F9A3F11CB828 CRC64;

Query Match 18.6%; Score 26; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHLP 15
| | | | |
Db 11 PHLP 14

RESULT 10
Q13665 ID Q13665 PRELIMINARY; PRT; 18 AA.
AC Q13665;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE CARDIAC MYOSIN BINDING PROTEIN (FRAGMENT).
GN MYBP-C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96083592; PubMed=7493025;

```
RA Watkins H., Conner D., Thierfelder L., Jarcho J.A., Macrae C.,
RA McKenna W.J., Maron B.J., Seidman J.G., Seidman C.E.;
RT "Mutations in the cardiac myosin binding protein-C gene on chromosome
RL 11 cause familial hypertrophic cardiomyopathy.";
RL Nat. Genet. 11:434-437(1995).
DR EMBL; S80805; AAB35661.1; -.
FT NON_TER 1
SQ SEQUENCE 18 AA; 1891 MW; 3EC842CE85A19C97 CRC64;

Query Match 18.6%; Score 26; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYQSRVT 10
   || | | |
Db 9 CGGIYVCRAT 18

RESULT 11
Q9TQZ6 PRELIMINARY; PRT; 20 AA.
AC Q9TQZ6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=96273610; PubMed=9690030;
RA Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT "Studies on the mechanism of early onset macular degeneration in
RT cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
RT of two proteins in the retina.";
RL Exp. Eye Res. 62:211-219(1996).
SQ SEQUENCE 20 AA; 2411 MW; 5F1A6AEB5918F777 CRC64;

Query Match 18.6%; Score 26; DB 6; Length 20;
Best Local Similarity 44.4%; Pred. No. 2.2e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ETYQSRVTH 11
   :|:| | |
Db 11 DTHKSEVAH 19

RESULT 12
Q52321 PRELIMINARY; PRT; 23 AA.
AC Q52321
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HSDS PROTEIN (FRAGMENT).
GN HSDS.
OS Escherichia coli.
OG Plasmid R124.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89137712; PubMed=3265670;
RA Price C., Bickle T.A.;
RT "Evolution of DNA sequence specificity in type I restriction
RT enzymes.";
```

```
RL Biochem. Soc. Trans. 16:942-943(1988).
DR EMBL; M27782; AAA56854.1; -.
KW Plasmid.
FT NON_TER 1
FT NON_TER 23
SQ SEQUENCE 23 AA; 2436 MW; 17351FBF990C1827 CRC64;

Query Match 18.6%; Score 26; DB 2; Length 23;
Best Local Similarity 54.5%; Pred. No. 2.6e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 13 HLPALMRSTT 23
   |||:| | |
Db 3 HPLSLPLSLT 13

RESULT 13
Q73337 PRELIMINARY; PRT; 24 AA.
AC Q73337
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NEF.
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HIV-1 C18MBC;
RA Deacon N.J., Tsykin A., Solomon A., Smith K., Ludford-Menting M.,
RA Hooker D.J., McPhee D.A., Greenway A.L., Ellett A., Chatfield C.,
RA Lawson V.A., Crowe S., Maerz A., Sonza S., Learmont J., Sullivan J.S.,
RA Cunningham A., Dwyer D., Dowton D., Mills J.;
RT "Genomic structure of an attenuated quasi species of HIV-1 from a
RT blood transfusion donor and recipients.";
RL Science 270:988-991(1995).
DR EMBL; U37270; AAC54551.1; -.
SQ SEQUENCE 24 AA; 2760 MW; 108D5DB2E4F7B73D CRC64;

Query Match 18.6%; Score 26; DB 15; Length 24;
Best Local Similarity 35.7%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GETYQSRVTHPLP 15
   |::| | | |
Db 3 GKWSESVVRRHVP 16

RESULT 14
O42416 PRELIMINARY; PRT; 19 AA.
AC O42416
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE THROMBOMUCIN (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97444372; PubMed=9298993;
RA McNaghy K.M., Pettersson I., Rossi F., Flamme I., Shevchenko A.,
RA Mann M., Graf T.;
RT "Thrombomucin, a novel cell surface protein that defines thrombocytes
RT and multipotent hematopoietic progenitors.";
```


RL J. Cell Biol. 138:1395-1407(1997).
DR EMBL; Y13977; CAA74310.1; -.
FT NON_TER 1 1
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2181 MW; 483C3DF97E13EC19 CRC64;

Query Match 18.2%; Score 25.5; DB 13; Length 19;
Best Local Similarity 42.9%; Pred. NO. 2.5e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

QY 11 HPHLP-----RALM 19
||| ||| |||::
Db 3 HPDLPGFDDGRAIL 16

RESULT 15

Q9BUH0 PRELIMINARY; PRT; 22 AA.
AC Q9BUH0;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 2.6 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC002629; AA002629.1; -.
KW Hypothetical protein.
SQ SEQUENCE 22 AA; 2645 MW; 60B9F471AB613874 CRC64;

Query Match 18.2%; Score 25.5; DB 4; Length 22;
Best Local Similarity 53.3%; Pred. NO. 3e+03;
Matches 8; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 8 RVTH-PHLPRALMRS 21
||| ||| | |:
Db 3 RVTWVPHLAVQLQRN 17

Search completed: August 26, 2002, 10:29:34
Job time: 649 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 26, 2002, 10:20:41 ; Search time 43.94 Seconds
(without alignments)
54.671 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPHLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 4981

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	29	20.6	20	A61093	glue protein - Cal
2	28	19.9	11	S78765	ribosomal protein
3	28	19.9	23	F61491	seed protein ws-6
4	27	19.1	15	C37765	hypothetical prote
5	27	19.1	17	PH1802	T cell receptor al
6	27	19.1	19	B53145	high conductance c
7	27	19.1	19	E49048	T-cell receptor be
8	27	19.1	20	A41439	acid ribonuclease
9	26.5	18.8	25	A58647	alphaA-conotoxin p
10	26	18.4	13	S32551	glutathione transf
11	26	18.4	14	PH1628	Ig H chain V-D-J r
12	26	18.4	15	I38336	hypothetical TEL/M
13	26	18.4	16	PH1634	Ig H chain V-D-J r
14	26	18.4	25	A60621	somatotropin - Atl
15	25	17.7	10	H37196	bradykinin-potenti
16	25	17.7	14	S62374	alpha-1-antichymot
17	25	17.7	15	PA0071	superoxide dismuta
18	25	17.7	15	I50503	agrin - electric r
19	25	17.7	17	S69164	ferredoxin al - Ja
20	25	17.7	21	B33600	glutamate--ammonia
21	25	17.7	22	PH1359	Ig heavy chain DJ
22	25	17.7	22	T01859	CXI la protein - c
23	25	17.7	23	S72535	probable acr-2 reg
24	25	17.7	24	B44379	omega-conotoxin SV
25	25	17.7	24	S29749	serum albumin - do
26	25	17.7	24	G85602	hypothetical prote
27	25	17.7	25	PH1686	Ig heavy chain V r
28	24.5	17.4	23	S23637	hypothetical prote
29	24	17.0	10	B37196	bradykinin-potenti

30	24	17.0	13	2	S47358	T-cell antigen rec
31	24	17.0	15	2	S42741	ubiquinol--cytochr
32	24	17.0	17	2	I49425	mitogen regulated
33	24	17.0	20	2	S23981	outer layer protei
34	24	17.0	22	2	S42567	cytochrome-b5 redu
35	24	17.0	23	2	S48156	alpha-amylase inhi
36	24	17.0	24	2	A25738	nicotinic acetylch
37	24	17.0	25	2	S03456	T-cell receptor al
38	23.5	16.7	14	2	A58963	alpha-conotoxin Cn
39	23	16.3	11	2	PT0301	Ig heavy chain CRD
40	23	16.3	14	1	NTKN1M	alpha-conotoxin MI
41	23	16.3	15	2	PA0009	seed storage prote
42	23	16.3	17	2	C85956	hypothetical prote
43	23	16.3	19	2	PA0012	superoxide dismuta
44	23	16.3	20	2	JA0142	proteinase inhibit
45	23	16.3	20	2	I67551	monocyte chemotact

ALIGNMENTS

RESULT 1

A61093
glue protein - California mussel (fragments)
N:Alternate names: adhesive polyphenolic protein
C:Species: Mytilus californianus (California mussel)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: A61093
R:Waite, J.H.
J. Comp. Physiol. B 156, 491-496, 1986
A:Title: Mussel glue from Mytilus californianus Conrad: a comparative study.
A:Reference number: A61093; MOID:86279063
A:Accession: A61093
A:Molecule type: protein
A:Residues: 1-20 <WAI>
A:Note: the amino terminal residue was tentatively identified as Ser
A:Note: 12-Ser was also found
C:Comment: this glue protein is a component of the adhesive plaque of the byssus, the
C:Keywords: hydroxyproline; tandem repeat
F:4,14/Modified site: 3-hydroxyproline (Pro) #status experimental
F:5,9,15,19/Modified site: 4-hydroxyproline (Pro) #status experimental
F:7,17/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental
F:13/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental

Query Match 20.6%; Score 29; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 2; Gaps 1;

QY 4 TYYSRVTHP--HLPK 16
|| :||: ||
DB 6 TVKPKITYPTTVKPK 20

RESULT 2

S78765
ribosomal protein MRP-S24, mitochondrial - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: S78765
R:Graack, H.R.
submitted to the Protein Sequence Database, July 1999
A:Reference number: S78760
A:Accession: S78765
A:Molecule type: protein
A:Residues: 1-11 <GRA>
C:Keywords: mitochondrial
F:1-11/Product: ribosomal protein MRP-S24 (fragment) #status experimental <MAT>
Query Match 19.9%; Score 28; DB 2; Length 11;
Best Local Similarity 40.0%; Pred. No. 6.1e+02;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPPLPKDIVR 20

Db 2 HVDVPKDLTK 11

RESULT 3

F61491

seed protein ws-6 - winged bean (fragment)

C:Species: Psophocarpus tetragonolobus (winged bean)

C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 03-May-1996

C:Accession: F61491

R:Hirano, H.

J. Protein Chem. 8, 115-130, 1989

A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-dimensional gels

A:Reference number: A61491; MUID:89351606

A:Accession: F61491

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <HIR>

C:Superfamily: pathogenesis-related protein

C:Keywords: seed

Query Match

Best Local Similarity 19.9%; Score 28; DB 2; Length 23;

Matches 6; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 4 TYYSRVTHPHLPKDIVRSI 22

Db 4 TYEDETTPVAPAILYKAI 22

RESULT 4

C37765

hypothetical protein (csmA 3' region) - Chloroflexus aurantiacus (fragment)

C:Species: Chloroflexus aurantiacus

C>Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993

C:Accession: C37765

R:Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.

J. Bacteriol. 172, 4497-4504, 1990

A:Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantiacus

A:Reference number: A37765; MUID:90330558

A:Accession: C37765

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-15 <THE>

A:Cross-references: GB:M33964

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 15;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 YYSRVTHP 12

Db 6 YYSHTTVP 13

RESULT 5

PH1802

T cell receptor alpha chain V region (clone 3PBL V alpha 24-8) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: PH1802

R:Porcellini, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.

J. Exp. Med. 178, 1-16, 1993

A:Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood

A:Reference number: PH1754; MUID:93301585

A:Accession: PH1802

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-17 <POR>

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 17;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGETVYS 7

Db 3 CGERPYS 9

RESULT 6

B53145

high conductance calcium-activated potassium channel, maxi-K channel - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 18-Jul-2001

C:Accession: B53145

R:Knaus, H.G.; Garcia-Calvo, M.; Kaczorowski, G.J.; Garcia, M.L.

J. Biol. Chem. 269, 3921-3924, 1994

A:Title: Subunit composition of the high conductance calcium-activated potassium channel

A:Reference number: A53145; MUID:94140798

A:Accession: B53145

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-19 <KNA>

A:Note: sequence extracted from NCBI backbone (NCBIP:144547)

C:Superfamily: fruit fly calcium-activated potassium channel slo

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 19;

Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 EYYSRVTHPH 13

Db 6 EYIQGSVLNPH 16

RESULT 7

E49048

T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997

C:Accession: E49048

R:Sioud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.

Eur. J. Immunol. 22, 2413-2418, 1992

A:Title: Limited heterogeneity of T cell receptor variable region gene usage in juven

A:Reference number: A49048; MUID:92387250

A:Accession: E49048

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-19 <SIO>

A:Experimental source: patient EV, IL-2R+ synovial T-cells

A:Note: sequence extracted from NCBI backbone (NCBIP:113267)

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 19;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGETVY 6

Db 9 CGRFY 14

RESULT 8

A41439

acid ribonuclease (EC 3.1.1.-) - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 21-Mar-1996

```

A;Experimental source: female
A;Accession: S32550
A:Molecule type: protein
A:Residues: 1-13 <SIN2>
A:Experimental source: male
C:Keywords: transferase

Query Match      18.4%   Score 26;   DB 2;   Length 13;
Best Local Similarity 80.0%;   Pred. No. 1.4e+03;
Matches      4;   Conservative      0;   Mismatches      0;   Indels      0;   Gaps      0;

Qy      8   RVTHP 12
         | : | |
Db      9   RLTHP 13

RESULT 11
PH1628
Iq H chain V-D-J region (clone B-less 151) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1628
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less
A:Reference number: PH1580; MUID:93301609
A:Accession: PH1628
A:Molecule type: DNA
A:Residues: 1-14 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match      18.4%   Score 26;   DB 2;   Length 14;
Best Local Similarity 44.4%;   Pred. No. 1.5e+03;
Matches      4;   Conservative      1;   Mismatches      4;   Indels      0;   Gaps      0;

Qy      1   CGETYYSRV 9
         | : | | :
Db      1   CARHYSNL 9

RESULT 12
I38336
Hypothetical TEL/MNL mutant fusion protein type I - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000
C:Accession: I38336
R:Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A.
Oncogene 10, 1511-1519, 1995
A:Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in
A:Reference number: I38031; MUID:95249265
A:Accession: I38336
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-15 <BUI>
A:Cross-references: EMBL:X85026; NID:g971473; PIDN:CAA59399.1; PID:g971474
C:Comment: This sequence is the chimeric product of a translocation mutation.
C:Genetics:
A:Gene: ETV6/MNL1; TEL/MNL1
A:Map position: 22q11/12p13
C:Keywords: fusion protein

Query Match      18.4%   Score 26;   DB 4;   Length 15;
Best Local Similarity 66.7%;   Pred. No. 1.7e+03;
Matches      4;   Conservative      1;   Mismatches      1;   Indels      0;   Gaps      0;

Qy      13  HLPKDI 18
         | | | :
Db      6  HLPHDL 11

```

RESULT 13

PHI634
Ig H chain V-D-J region (clone B-less 213) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PHI634
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A:Reference number: PHI580; MUID:93301609
A:Accession: PHI634
A:Molecule type: DNA
A:Residues: 1-16 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 18.4%; Score 26; DB 2; Length 16;

Best Local Similarity 41.7%; Pred. No. 1.8e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

QY 3 ETYYSRVTHPHL 14

Db :||| ||:

DSYIS----PHI 11

RESULT 14

A60621
Somatotropin - Atlantic salmon (fragment)
N:Alternate names: growth hormone
C:Species: Salmo salar (Atlantic salmon)
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C:Accession: A60621
R:Skibell, V.; Andersen, O.; Gautvik, K.M.
Gen. Comp. Endocrinol. 80, 333-344, 1990
A:Title: Purification and characterization of Atlantic salmon growth hormone and evidence of its biological activity
A:Reference number: A60621; MUID:91146880
A:Accession: A60621
A:Molecule type: protein
A:Residues: 1-25 <SKI>
A:Note: this protein displayed charge heterogeneity from variable levels of phosphorylation
C:Superfamily: prolactin
C:Keywords: hormone; phosphoprotein; pituitary

Query Match 18.4%; Score 26; DB 2; Length 25;

Best Local Similarity 62.5%; Pred. No. 2.8e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 SRVTHPHL 14

Db :||| ||:

NRVOHLHL 19

RESULT 15

H37196
bradykinin-potentiating peptide 8 - island jararaca
C:Species: Bothrops insularis (island jararaca)
C:Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994
C:Accession: H37196
R:Cintrá, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides from Bothrops insularis
A:Reference number: A37196; MUID:90351557
A:Accession: H37196
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <GIN>
C:Keywords: pyroglutamic acid
F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 17.7%; Score 25; DB 2; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPHLP 15

Db ||:|

5 HPNIP 9

Search completed: August 26, 2002, 10:20:41
Job time: 347 sec

44-38861-1054

THIS PAGE BLANK (USPTO)

PPK4_PERAM
ID PPK4_PERAM STANDARD; PRT; 12 AA.
AC P82619;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pyrokinin-4 (Pea-PK-4) (YXPRL-amide).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE-Retrocerebral complex;
RX MEDLINE=99212469; PubMed=10196736;
RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;
RT "Differential distribution of pyrokinin-isoforms in cerebral and abdominal neurohemal organs of the American cockroach.";
RL Insect Biochem. Mol. Biol. 29:139-144(1999).
RN [2]
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of the American cockroach.";
RL J. Comp. Neurol. 419:352-363(2000).
CC -1- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY (MYOTROPIC ACTIVITY).
CC -1- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC -1- MASS SPECTROMETRY: MW=1147.9; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 12 12
FT MOD_RES 12 12
SQ SEQUENCE 12 AA; 1449 MW; FA7A3049FF42CA1 CRC64;

Query Match 19.1%; Score 27; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 HLPKDI 18
DB 2 HLPHDV 7

RESULT 3
CX44_CONPU STANDARD; PRT; 25 AA.
AC P55963;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-A conotoxin PIVA.
OS Conus purpurascens (Purple cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=41690;
RN [1]
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=95403432; PubMed=7673220;
RA Hopkins C., Grille M., Miller C., Shon K.-J., Cruz L.J., Gray W.R., Dykert J., Rivier J., Yoshikami D., Olivera B.M.;
RT "A new family of Conus peptides targeted to the nicotinic acetylcholine receptor.";
RL J. Biol. Chem. 270:22361-22367(1995).
RN [2]
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=97200721; PubMed=9048550;
RA Han K.-H., Hwang K.-J., Kim S.-M., Kim S.-K., Gray W.R., Olivera B.M., Rivier J., Shon K.-J.;
RT "NMR structure determination of a novel conotoxin, [Pro 7,13] alpha

RT A-conotoxin PIVA.";
RL Biochemistry 36:1669-1677(1997).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS INHIBIT THEM.
CC -1- SUBCELLULAR LOCATION: Secreted.
DR PDB; 1P1P; 07-JUL-97.
CC -1- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation; Hydroxylation; Venom; 3D-structure.
FT DISULFID 2 16
FT DISULFID 3 11
FT DISULFID 14 23
FT MOD_RES 7 7
FT MOD_RES 13 13
FT MOD_RES 20 20
FT MOD_RES 25 25
SQ SEQUENCE 25 AA; 2608 MW; 9E2147898D697640 CRC64;

Query Match 18.8%; Score 26.5; DB 1; Length 25;
Best Local Similarity 41.2%; Pred. No. 8.1e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 CGETYSRVTHPHLPK 17
DB 3 CG-SYPNAACHPCSKD 18

RESULT 4
UN05_PINPS STANDARD; PRT; 20 AA.
AC P81674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Unknown protein from 2D-page of needles (N147) (Fragments).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A., Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN PROTEIN IS: 5.6, ITS MW IS: 36 kDa.
FT NON_TER 1 1
FT NON_TER 11 12
FT NON_TER 13 13
FT NON_TER 14 14
FT NON_TER 15 15
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2438 MW; 9F4E4678E086C298 CRC64;

Query Match 18.4%; Score 26; DB 1; Length 20;
Best Local Similarity 31.2%; Pred. No. 7.6e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 1 CGETYSRVTHPHLPK 16
DB 9 CDRDFY----RPELPR 20

RESULT 5
BPP8_BOTIN STANDARD; PRT; 10 AA.
ID BPP8_BOTIN

AC P30426;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Bradykinin-potentiating peptide S5,1 (Angiotensin-converting
 enzyme inhibitor).
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8723;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Venom;
 RC MEDLINE=90351557; PubMed=2386615;
 RX Cintra A.C.O., Vieira C.A., Giglio J.R.;
 RT "Primary structure and biological activity of bradykinin potentiating
 RT peptides from Bothrops insularis snake venom.";
 RL J. Protein Chem. 9:221-227(1990).
 CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
 CC PIR; H37196; H37196.
 DR Hypotensive agent; Venom.
 FT MOD_RES 1
 FT SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;
 SQ
 Query Match 17.7%; Score 25; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 5.2e+02;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 11 HPHP 15
 DB 5 HPNP 9
 RESULT 6
 MCA2_RHOOP STANDARD; PRT; 15 AA.
 ID MCA2_RHOOP
 AC P56870;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative maleylacetate reductase II (EC 1.3.1.32) (Fragment).
 OS Rhodococcus opacus (Nocardia opaca).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
 OX NCBI_TaxID=37919;
 RN [1]
 RN SEQUENCE.
 RP STRAIN=ICP;
 RC MEDLINE=98324954; PubMed=9657989;
 RX Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;
 RT "Characterization of a maleylacetate reductase encoding region from
 RT Rhodococcus opacus ICP.";
 RL J. Bacteriol. 180:3503-3508(1998).
 CC -1- CATALYTIC ACTIVITY: 3-oxoadipate + NAD(P)(+) = 2-maleylacetate +
 CC NAD(P)H.
 CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).
 CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC
 CC AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL
 CC PRODUCTS AND AS INDUSTRIAL EFFLUENT.
 CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE
 CC FAMILY.
 CC InterPro; IPR001670; Fe-ADH.
 DR PROSITE; PS00913; ADH_IRON_1; PARTIAL.
 DR PROSITE; PS00060; ADH_IRON_2; PARTIAL.
 KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
 FT NON_TER 15
 FT SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 17.7%; Score 25; DB 1; Length 15;
 Best Local Similarity 45.5%; Pred. No. 8e+02;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 8 RVTPHPHPKDI 18
 DB 2 RFHENLPQRI 12
 RESULT 7
 BPP2_BOTIN STANDARD; PRT; 10 AA.
 ID BPP2_BOTIN
 AC P30422;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Bradykinin-potentiating peptide S4,3,1 (10C) (Angiotensin-converting
 enzyme inhibitor).
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 OX NCBI_TaxID=8723;
 RN [1]
 RN SEQUENCE.
 RP TISSUE=Venom;
 RC MEDLINE=90351557; PubMed=2386615;
 RX Cintra A.C.O., Vieira C.A., Giglio J.R.;
 RT "Primary structure and biological activity of bradykinin potentiating
 RT peptides from Bothrops insularis snake venom.";
 RL J. Protein Chem. 9:221-227(1990).
 CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
 CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
 CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
 CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
 CC PIR; B37196; B37196.
 DR Hypotensive agent; Venom.
 FT MOD_RES 1
 FT SEQUENCE 10 AA; 1213 MW; 30C53546C761F773 CRC64;
 SQ
 Query Match 17.0%; Score 24; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 7.4e+02;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 11 HPHP 15
 DB 5 HPQIP 9
 RESULT 8
 AU11_LITRA STANDARD; PRT; 13 AA.
 ID AU11_LITRA
 AC P82386;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Aurein 1.1.
 OS Litoria raniformis (Southern bell frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hyllidae;
 OC Litoria.
 OX NCBI_TaxID=116057;
 RN [1]
 RN SEQUENCE, AND FUNCTION.
 RC TISSUE=Granular dorsal gland;
 RX MEDLINE=20408845; PubMed=10951191;
 RA Rozek T., Wegener K.L., Bowie J.H., Oliver I.N., Carver J.A.,
 RA Wallace J.C., Tyler M.J.;
 RT "The antibiotic and anticancer active aurein peptides from the
 RT Australian bell frogs Litoria aurea and Litoria raniformis the
 RT solution structure of aurein 1.2.";

```

RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
CC L.INNOCUA AND S.UBERIS. PROBABLY ACTS BY DISTURBING MEMBRANE
CC FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Antibiotic: Amidation.
FT MOD_RES 13 13
SQ SEQUENCE 13 AA; 1447 MW; 173CB99DFBC83330 CRC64;

Query Match 17.0%; Score 24; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 9.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 17 DIVRSIAK 24
Db 4 DIKKIAE 11

RESULT 9
AUI2_LITRA
ID AUI2_LITRA STANDARD; PRT; 13 AA.
AC P82387;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Aurein 1.2.
OS Litoria raniformis (Southern bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoldea; Hylidae;
OC Litoria.
OX NCBI_TaxID=116057;
RN [1]
RP SEQUENCE, FUNCTION, AND STRUCTURE BY NMR.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2.";
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST B.CEREUS, L.LACTIS,
CC L.INNOCUA, M.LUTEUS, P.MULFOCIDA, S.AUREUS, S.EPIDERMIDIS AND
CC S.UBERIS. PROBABLY ACTS BY DISTURBING MEMBRANE FUNCTIONS WITH ITS
CC AMPHIPATHIC STRUCTURE. SHOWS ANTICANCER ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT MOD_RES 13 13
SQ SEQUENCE 13 AA; 1481 MW; 1EACB99DFBC83330 CRC64;

Query Match 17.0%; Score 24; DB 1; Length 13;
Best Local Similarity 50.0%; Pred. No. 9.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 17 DIVRSIAK 24
Db 4 DIKKIAE 11

RESULT 10
AUI3_LITRA
ID AUI3_LITRA STANDARD; PRT; 17 AA.
AC P82394;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Aurein 3.1 [Contains: Aurein 3.1.1; Aurein 3.1.2].
OS Litoria raniformis (Southern bell frog), and
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoldea; Hylidae;

```

```

OC Litoria.
OX NCBI_TaxID=116057, 8371;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2.";
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST L.LACTIS,
CC L.INNOCUA, M.LUTEUS, S.AUREUS, S.EPIDERMIDIS AND S.UBERIS.
CC PROBABLY ACTS BY DISTURBING MEMBRANE FUNCTIONS WITH ITS
CC AMPHIPATHIC STRUCTURE. AUREIN 3.1 SHOWS ANTICANCER ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT PEPTIDE 1 17 AUREIN 3.1.1.
FT PEPTIDE 1 14 AUREIN 3.1.1.1.
FT PEPTIDE 3 17 AUREIN 3.1.1.2.
FT MOD_RES 17 17 AMIDATION.
SQ SEQUENCE 17 AA; 1739 MW; F6AC17980428EAFD CRC64;

Query Match 17.0%; Score 24; DB 1; Length 17;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 DIVRSIA 23
Db 4 DIVKKIA 10

RESULT 11
AUI32_LITRA
ID AUI32_LITRA STANDARD; PRT; 17 AA.
AC P82395;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aurein 3.2.
OS Litoria raniformis (Southern bell frog), and
OS Litoria aurea (Green and golden bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoldea; Hylidae;
OC Litoria.
OX NCBI_TaxID=116057, 8371;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2.";
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST L.LACTIS, L.INNOCUA,
CC M.LUTEUS, S.AUREUS, S.EPIDERMIDIS AND S.UBERIS. PROBABLY ACTS BY
CC DISTURBING MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC SHOWS ANTICANCER ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT MOD_RES 17 17 AMIDATION.
SQ SEQUENCE 17 AA; 1769 MW; F6AC0A580428EAFD CRC64;

Query Match 17.0%; Score 24; DB 1; Length 17;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 DIVRSIA 23

```

```
Db 4 DIVKIA 10
||||: ||
SEQUENCE 20 AA; 2325 MW; 45FC7989AB7527C7 CRC64;

RESULT 12
AU33_LITRA
ID AU33_LITRA STANDARD; PRT; 17 AA.
AC P82396;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Aurein 3.3 [Contains: Aurein 3.3.1].
OS Litoria raniformis (Southern bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.
OX NCBI_TaxID=116057;
RN [1]
RP SEQUENCE, AND FUNCTION.
RC TISSUE=Granular dorsal gland;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT australian bell frogs Litoria aurea and Litoria raniformis the
RT solution structure of aurein 1.2";
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: ANTIMICROBIAL ACTIVITY AGAINST L.LACTIS, M.LUTEUS,
CC PP.MULTOCIDIA, S.AUREUS, S.EPIDERMIDIS AND S.UBERIS. PROBABLY ACTS
CC BY DISTURBING MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC SHOWS ANTICANCER ACTIVITY.
CC -!- SUBCELLULAR LOCATION: SECRETED.
KW Amidation; Antibiotic.
FT PEPTIDE 1 16 AUREIN 3.3.
FT PEPTIDE 3 16 AUREIN 3.3.1.
FT MOD_RES 17 17 AMIDATION.
SQ SEQUENCE 17 AA; 1797 MW; F6AC0A471428EAFD CRC64;

Query Match 17.0%; Score 24; DB 1; Length 17;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 17 DIVRSTA 23
Db 4 DIVKIA 10
||||: ||

RESULT 13
VMO2_CHICK
ID VMO2_CHICK STANDARD; PRT; 20 AA.
AC Q9PS49;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vitelline membrane outer layer protein II (VMO-II) (VMOII) (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Egg yolk;
RX MEDLINE=92392273; PubMed=1520265;
RA Kido S., Morimoto A., Kim F., Doi Y.;
RT "Isolation of a novel protein from the outer layer of the vitelline
RT membrane.";
RL Biochem. J. 286:17-22(1992).
CC -!- FUNCTION: EXACT FUNCTION NOT KNOWN, COMPONENT OF THE OUTER
CC MEMBRANE OF THE VITELLINE LAYER OF THE EGG.
CC -!- PTM: ALL CYSTEINE RESIDUES OF THE MATURE PROTEIN ARE INVOLVED IN
CC DISULFIDE BONDS.
```

```
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2325 MW; 45FC7989AB7527C7 CRC64;

Query Match 17.0%; Score 24; DB 1; Length 20;
Best Local Similarity 44.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 14 LPRDIVRSI 22
Db 1 LPRDTSRXV 9
||||: ||
SEQUENCE 20 AA; 2325 MW; 45FC7989AB7527C7 CRC64;

RESULT 14
CXAL_CONCN
ID CXAL_CONCN STANDARD; PRT; 14 AA.
AC P56973;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alpha-conotoxin Cn1A [Contains: Alpha-conotoxin Cn1B].
OS Conus consors (Singed cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101297;
RN [1]
RP SEQUENCE, SYNTHESIS, CHARACTERIZATION, AND STRUCTURE BY NMR.
RC TISSUE=Venom;
RX MEDLINE=99255390; PubMed=10320362;
RA Favreau P., Krimm I., le Gall F., Bobenrieth M.J., Lamthanh H.,
RA Bouet F., Servent D., Molgo J., Menez A., Letourneux Y.,
RA Lancelin J.-M.;
RT "Biochemical characterization and nuclear magnetic resonance
RT structure of novel alpha-conotoxins isolated from the venom of Conus
RT consors.";
RL Biochemistry 38:6317-6326(1999).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM. THIS PEPTIDE SEEMS TO BE A POTENT AND SELECTIVE
CC BLOCKER OF MUSCULAR SUBTYPE OF NACHR.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
DR PDB: 1B45; 09-JUL-99.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; 3D-structure.
FT PEPTIDE 1 14 ALPHA-CONOTOXIN CN1A.
FT PEPTIDE 3 14 ALPHA-CONOTOXIN CN1B.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14 AMIDATION.
SQ SEQUENCE 14 AA; 1548 MW; DEEE91969BF5E5BD CRC64;

Query Match 16.7%; Score 23.5; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGETYYYS 7
Db 8 CGK-YYs 13
||||: ||

RESULT 15
CXAL_CONMA
ID CXAL_CONMA STANDARD; PRT; 14 AA.
AC P01521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-2001 (Rel. 40, Last annotation update)
DE Alpha-conotoxin MI (MI).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
```

```

OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE.
RX MEDLINE=83073458; PubMed=7149738;
RA McIntosh J.M.; Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RT "Isolation and structure of a peptide toxin from the marine snail
RL Conus magus.";
RL Arch. Biochem. Biophys. 218:329-334(1982).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin MI. Disulfide bonding and conformational states.";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: BELONGS TO THE ALPHA-TYPE CONOTOXIN FAMILY.
DR PIR; A01784; NTKNIM.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14 AMIDATION.
SQ SEQUENCE 14 AA; 1499 MW; DEEE91898BF5E5BD CRC64;

Query Match 16.3%; Score 23; DB 1; Length 14;
Best Local Similarity 60.0%; Pred. No. 1.5e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 CGEY 5
Db 8 CCKNY 12

```

Search completed: August 26, 2002, 10:30:17
Job time: 637 sec